Evolution Equations and Control Theory
Vol. 25, 2026, pp. 183-233 (July 29, 2026)
doi:10.3934/eect.2026088

GLOBAL-IN-TIME OPTIMAL CONTROL OF STOCHASTIC THIRD-GRADE FLUIDS WITH ADDITIVE NOISE

Kush Kinra*[1,2] and Fernanda Cipriano[3]

[1]Center for Mathematics and Applications (NOVA Math),
NOVA School of Science and Technology (NOVA FCT), Portugal

[2]Chair for Dynamics, Control, Machine Learning and Numerics, Department of Mathematics,
Friedrich-Alexander-Universitat Erlangen-Nürnberg, 91058 Erlangen, Germany

[3]Center for Mathematics and Applications (NOVA Math) and Department of Mathematics,
NOVA School of Science and Technology (NOVA FCT), Portugal

(Communicated by Enrico Priola)

Abstract. In this article, we address the velocity tracking control problem for a class of stochastic non-Newtonian fluids. More precisely, we consider the stochastic third-grade fluid equation perturbed by infinite-dimensional additive white noise and defined on the two-dimensional torus $\mathbb{T}^2$. The control acts as a distributed random external force. Taking an *infinite-dimensional Ornstein-Uhlenbeck process*, the stochastic system is converted into an equivalent pathwise deterministic one, which allows to show the well-posedness of the original stochastic system globally in time. The state being a stochastic process with sample paths in $\mathrm{L}^\infty(0,T;\mathbb{H}^3(\mathbb{T}^2))$ and finite moments can be controlled in an optimal way. Namely, we establish the existence and uniqueness of solutions to the corresponding linearized state and adjoint equations. Furthermore, we derive an appropriate stability result for the state equation and verify that the Gâteaux derivative of the control-to-state mapping coincides with the solution of the linearized state equation. Finally, we prove the existence of an optimal solution and establish the first-order necessary optimality conditions.

1. **Introduction.** The optimal control of the evolution of fluid flows is a significant mathematical challenge with great implications in real-world applications. This is particularly true for non-Newtonian fluids of differential type, which exhibit complex rheological behaviors that depend on the rate of deformation. Unlike Newtonian fluids, which satisfy the linear Newtonian viscosity law, these fluids display nonlinear stress-strain relationships, leading to unique phenomena such as shear-thinning, shear-thickening, and viscoelastic effects (see, for example, [14] and

2020 *Mathematics Subject Classification.* Primary: 35R60, 49K20, 76A05; Secondary: 76D55, 60H15.

Key words and phrases. Stochastic third-grade fluids, optimal control, necessary optimality condition, infinite-dimensional Wiener process.

This work is funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UID/00297/2025 (https://doi.org/10.54499/UID/00297/2025) and UID/PRR/00297/2025 (https://doi.org/10.54499/UID/PRR/00297/2025) (Center for Mathematics and Applications).

*Corresponding author: Kush Kinra.

references therein). These properties are critical in various industrial and technological processes, including the manufacturing of polymers, the processing of food products, biomedical applications such as blood flow modeling and many others. The ability to control and optimize these flows is essential for improving efficiency, ensuring process stability, and enhancing product quality. However, due to the strong nonlinearity and complexity of their governing equations, the mathematical treatment of their optimal control is not an easy issue, requiring involving analysis and sophisticated mathematical tools. Non-Newtonian fluids of differential type constitute a hierarchy of complex fluids, the rheological relation of fluids of grade n, $n>1$, is nonlinear, and its constitutive law reads

$$\mathbb{S} = -p\mathrm{I} + \mathrm{F}(\mathrm{A}_1(\boldsymbol{v}), \cdots, \mathrm{A}_n(\boldsymbol{v})),$$

where $\mathbb{S}$ is the Cauchy stress tensor and F is an isotropic polynomial function of degree n subject to the usual requirement of material frame indifference, $\boldsymbol{v}$ is fluid's velocity field and $\mathrm{A}_n, n \geq 1$, are the Rivlin-Ericksen kinematic tensors (see [23]), which are defined by

$$\begin{cases} \mathrm{A}_1(\boldsymbol{v}) = \nabla \boldsymbol{v} + \nabla \boldsymbol{v}^T, \\ \mathrm{A}_n(\boldsymbol{v}) = \dfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{A}_{n-1}(\boldsymbol{v}) + \mathrm{A}_{n-1}(\boldsymbol{v})(\nabla \boldsymbol{v}) + (\nabla \boldsymbol{v})^T \mathrm{A}_{n-1}(\boldsymbol{v}), \quad n = 2, 3, \cdots, \end{cases}$$

where $\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + (\boldsymbol{v} \cdot \nabla)$ which is also known as material derivative.

This work is devoted to third-grade fluids, which are characterized by the following constitutive law

$$\begin{aligned} \mathbb{S} = &-p\mathrm{I} + \nu \mathrm{A}_1(\boldsymbol{v}) + \alpha_1 \mathrm{A}_2(\boldsymbol{v}) + \alpha_2 \mathrm{A}_1^2(\boldsymbol{v}) + \beta_1 \mathrm{A}_3(\boldsymbol{v}) \\ &+ \beta_2(\mathrm{A}_1(\boldsymbol{v})\mathrm{A}_2(\boldsymbol{v}) + \mathrm{A}_2(\boldsymbol{v})\mathrm{A}_1(\boldsymbol{v})) + \beta_3 \mathrm{Tr}(\mathrm{A}_1^2(\boldsymbol{v}))\mathrm{A}_1(\boldsymbol{v}), \end{aligned}$$

where ν is the viscosity, and $\{\alpha_i\}_{i=1,2}$ and $\{\beta_i\}_{i=1,2,3}$ are material moduli. The momentum equations are given by

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = \mathrm{div}(\mathbb{S}).$$

The compatibility with thermodynamic laws imposes the following restriction on the parameters (see [14])

$$\nu \geq 0, \quad \alpha_1 \geq 0, \quad |\alpha_1 + \alpha_2| \leq \sqrt{24\nu\beta}, \quad \beta_1 = \beta_2 = 0, \, \beta_3 = \beta \geq 0. \tag{1}$$

Let us notice that $\beta = 0$ corresponds to the second grade fluids, and $\alpha_1 = \alpha_2 = 0$ and β=0 gives the Navier-Stokes equations. To better understand and describe the properties of various nanofluids, numerous simulation studies have been carried out using third-grade fluid models (see [16, 20, 22] and references therein). Nanofluids are engineered colloidal suspensions of nanoparticles-typically made of metals, oxides, carbides, or carbon nanotubes-dispersed in a base fluid such as water, ethylene glycol, or oil. These nanofluids exhibit higher thermal conductivity than the base fluid and hold significant promise for a wide range of technological applications, including heat transfer systems, microelectronics, fuel cells, pharmaceutical processes, hybrid-powered engines, engine cooling, and vehicle thermal management. Therefore, investigating the kinematic properties of third-grade fluids holds significant importance in various applications. The governing equations for third-grade fluids are expressed as:

$$\begin{cases}\partial_t(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v})-\nu\Delta\boldsymbol{v}+(\boldsymbol{v}\cdot\nabla)(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v})+\displaystyle\sum_{j=1}^{2}[\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v}]^j\nabla\boldsymbol{v}^j \\ -(\alpha_1+\alpha_2)\mathrm{div}([\mathrm{A}(\boldsymbol{v})]^2)-\beta\mathrm{div}[\mathrm{Tr}(\mathrm{A}(\boldsymbol{v})[\mathrm{A}(\boldsymbol{v})]^T)\mathrm{A}(\boldsymbol{v})]=-\nabla p+\boldsymbol{f}, \\ \mathrm{div}\ \boldsymbol{v}=0,\end{cases} \tag{2}$$

where $\boldsymbol{v}(x,t):\mathbb{T}^2\times[0,\infty)\to\mathbb{R}^2$, $p(x,t):\mathbb{T}^2\times[0,\infty)\to\mathbb{R}$ and $\boldsymbol{f}(x,t):\mathbb{T}^2\times[0,\infty)\to\mathbb{R}^2$ represent velocity field, pressure field and external forcing, respectively, at time t and position x, $[\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v}]^j$ and $\boldsymbol{v}^j$ denotes the j^{th} component of these vectors, $\mathrm{A}(\boldsymbol{v}):=\mathrm{A}_1(\boldsymbol{v})$, $\nu>0$ denotes the viscosity of the fluid and $\alpha_1>0$, $\alpha_2\in\mathbb{R}$, $\beta>0$ are material moduli verifying (1). In this article, we are interested in the stochastic version of the system (2) which is given by

$$\begin{cases}\mathrm{d}(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v})+\big\{-\nu\Delta\boldsymbol{v}+(\boldsymbol{v}\cdot\nabla)(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v})+\displaystyle\sum_{j=1}^{2}[\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v}]^j\nabla\boldsymbol{v}^j \\ -(\alpha_1+\alpha_2)\mathrm{div}([\mathrm{A}(\boldsymbol{v})]^2)-\beta\mathrm{div}[\mathrm{Tr}(\mathrm{A}(\boldsymbol{v})[\mathrm{A}(\boldsymbol{v})]^T)\mathrm{A}(\boldsymbol{v})]\big\}\mathrm{d}t \\ =(-\nabla p+\boldsymbol{f})\,\mathrm{d}t+\mathrm{dW}, & \text{in } \mathbb{T}^2\times(0,\infty), \\ \mathrm{div}\ \boldsymbol{v}=0, & \text{in } \mathbb{T}^2\times[0,\infty), \\ \boldsymbol{v}(x,0)=\boldsymbol{v}_0(x), & \text{in } \mathbb{T}^2,\end{cases} \tag{3}$$

where $\mathrm{W}(\cdot)$ is a GG^*-Wiener process defined in some given probability space $(\Omega,\mathscr{F},\mathbb{P})$, G being a Hilbert-Schmidt operator from $\mathbb{U}$ to $\mathbb{H}$. Here $\mathbb{U}$ denotes some Hilbert space and $\mathbb{H}$ is the space of functions in $\mathbb{L}^2(\mathbb{T}^2)$ which are mean-zero and divergence-free. For $\mathrm{L}>0$, we identify the two-dimensional torus $\mathbb{T}^2=\left(\frac{\mathbb{R}}{\mathrm{L}\mathbb{Z}}\right)^2$, with the rectangle $[0,\mathrm{L}]^2$ with periodic boundary conditions (see [15, Chapter 3]), namely $\boldsymbol{v}(\cdot,\cdot)$, $p(\cdot,\cdot)$ and $\boldsymbol{f}(\cdot,\cdot)$ satisfy the following periodic conditions:

$$\boldsymbol{v}(x+\mathrm{L}e_i,\cdot)=\boldsymbol{v}(x,\cdot),\ p(x+\mathrm{L}e_i,\cdot)=p(x,\cdot) \text{ and } \boldsymbol{f}(x+\mathrm{L}e_i,\cdot)=\boldsymbol{f}(x,\cdot), \tag{4}$$

$x\in\mathbb{R}^2$, $i\in\{1,2\}$, where $\{e_1,e_2\}$ is the canonical basis of $\mathbb{R}^2$.

Remark 1.1. We denote by $\mathrm{W}(\cdot)$ a GG^*-Wiener process on $\mathbb{H}$, that is, an $\mathbb{H}$-valued Q-Wiener process with covariance operator

$$Q=GG^*.$$

Equivalently, $\mathrm{W}(\cdot)$ is a centered Gaussian process satisfying

$$\mathbb{E}\big[\langle\mathrm{W}(t),h_1\rangle_{\mathbb{H}}\langle\mathrm{W}(s),h_2\rangle_{\mathbb{H}}\big]=(t\wedge s)\langle Qh_1,h_2\rangle_{\mathbb{H}},\quad h_1,h_2\in\mathbb{H}.$$

A GG^*-Wiener process can be represented as $\mathrm{W}(t)=G\widetilde{\mathrm{W}}(t)$, where $\widetilde{\mathrm{W}}(t)$ is a cylindrical Wiener process on $\mathbb{U}$. We refer the reader to [13, 17] for further details.

The main objective of this article is to investigate an optimal control problem associated with the stochastic state equation (3) subject to the periodic boundary conditions (4). The resolution of the control problem involves three key steps:

i) Global-in-time solvability of system (3) with an appropriate regularity to ensure the well-posedness of linearized state equation.
ii) Analyzing the Gâteaux derivative of the control-to-state mapping, which is characterized by the unique solution of the linearized state equation and addressing the well-posedness of the adjoint equation.

iii) Deriving the connection between the solution of the linearized equation and the adjoint state solution, which enables the formulation of the necessary first-order necessary optimality condition in terms of the adjoint state.

Without exhaustiveness, the third-grade fluid equations with Dirichlet boundary condition were studied in [1, 26]. Later on in [6, 5], the authors proved the existence of a global solution with $\mathbb{H}^2$ initial data and Navier-slip boundary condition in 2D as well as 3D, and demonstrated that uniqueness holds in 2D. The authors of [10] extended the later deterministic findings to 2D stochastic models. In [30], for stochastic third-grade fluids equations perturbed by multiplicative white noise, the authors proved that the local (in time) solution exists and is unique. Such solution corresponds to an adapted stochastic process with sample paths in $\mathbb{H}^3$, defined up to a specific positive stopping time. We also mention some works discussing optimal control problem for stochastic Navier-Stokes equations, see [3, 8, 11, 12, 18, 27] etc. Compared to the Navier–Stokes equations, the primary challenge in third-grade fluid equations arises from the presence of third-order derivatives in the nonlinear term, which significantly complicates the analysis. An optimal control problem for stochastic second-grade fluids has been addressed in [7]. In the framework of multiplicative noise, it is worth mentioning that the derivation of first-order necessary optimality conditions for the Navier–Stokes equations, as well as for second-grade fluid equations, is possible only under suitable restrictions on the model parameters (see [3, 7]). The optimal control of a generic stochastic fluid flow is not an easy issue.

This work aims to solve the optimal control problem for stochastic third-grade fluid equations (3) with additive noise, without additional restrictions on the physical parameters, and over a prescribed time interval. In order to solve the stochastic optimal control problem for the random vector field $\boldsymbol{v}$ governed by the stochastic third-grade fluid equations (3), we first prove that $\boldsymbol{v}$ satisfies the following key exponential integrability condition

$$\mathbb{E}\left[\exp\left\{\hat{c}\int_0^T \|\boldsymbol{v}(s)\|_{\mathbb{H}^3}^2 \mathrm{d}s\right\}\right] < +\infty, \quad \text{for } \hat{c} > 0. \tag{5}$$

To handle the additive white noise, we decompose the stochastic system (3) into two parts: one part is a linear stochastic differential equation involving the white noise, while the other one corresponds to a random partial differential equation. Under suitable assumptions on the white noise W, we establish the aforementioned regularity of the solution to stochastic system (3), which plays a key role in proving the well-posedness of the linearized state equation and demonstrating that the Gâteaux derivative of the cost functional coincides with the solution of the linearized state equation.

To the best of our knowledge, this is the first work, which addresses the global-in-time optimal control problem for stochastic third-grade fluid models. Recently, the authors in [29] considered the system (2) perturbed by a multiplicative noise and addressed the corresponding optimal control problem. However, due to the lack of global-in-time $\mathbb{H}^3$-regularity of the sample paths, the authors in [29] were able to control the state just locally-in-time. Namely, the cost functional was defined locally by taking an appropriate stopping time, and the stopping time plays a major role at each step of the analysis. In contrast to [29], the present work tackles the global-in-time optimal control problem associated with system (3). We also mention the work [28] where authors considered optimal control problem for deterministic third-grade fluid equations (2) supplemented by Navier-slip boundary condition.

Remark 1.2. It is worth mentioning that the extension of the present work to the three-dimensional stochastic third-grade fluid equations remains highly challenging. Although global weak solutions can be constructed in three dimensions, the uniqueness of solutions and their continuous dependence on the initial data and forcing terms are not known in general. These properties play a fundamental role in optimal control theory, since they are closely related to the well-posedness and stability of the control-to-state mapping.

As a consequence, the development of a global-in-time optimal control theory for three-dimensional third-grade fluids remains an open problem even in the deterministic setting. The stochastic framework considered here would require overcoming these unresolved deterministic difficulties in addition to establishing suitable probabilistic estimates, stability properties, and sensitivity results for the state equation.

Moreover, the derivation of first-order necessary optimality conditions typically requires sensitivity analysis of the state equation and the solvability of an associated adjoint system. The absence of uniqueness and stability estimates, together with the stronger nonlinear effects arising from the third-grade constitutive terms and the presence of stochastic forcing, makes the analysis of the corresponding three-dimensional velocity tracking problem considerably more involved.

Therefore, the analysis of global-in-time velocity tracking optimal control problems for three-dimensional stochastic third-grade fluids appears to require substantially new ideas and techniques.

The current paper's plan is as follows. We begin Section 2 by introducing the appropriate functional spaces. Then, we define linear and nonlinear operators. In addition, we present a linear stochastic equation along with its solvability result, which enables us to transform the stochastic equation (3) into an equivalent pathwise deterministic system (see (17) below). We conclude the section by presenting the main result of the article. In Section 3, we prove the existence and uniqueness of global-in-time solution of system (3) with appropriate regularity which is required to investigate our main result. Section 4 is devoted to the well-posedness of linearized state equation. In Section 5, we address the Gâteaux differentiability of the control-to-state mapping. Section 6 discusses the well-posedness of the adjoint system corresponding to the linearized state equation. The existence of a solution to the control problem is shown in the final section by establishing a duality property between the solutions of the linearized equation and the adjoint equation. Ultimately, by utilizing the duality relation, we further show that the solution to the control problem satisfies the first-order necessary optimality conditions.

2. Notations and auxiliary results.

The goal of this section is to provide necessary notations, function spaces, auxiliary results and the main goal of this article.

2.1. Function spaces.

Let us consider $m \in \mathbb{N}$ and $p > 0$. The usual norms for scalar (resp. vector) valued functions in the classical Lebesgue and Sobolev spaces $\mathrm{L}^p(\mathbb{T}^2)$ (resp. $\mathbb{L}^p(\mathbb{T}^2)$), $\mathrm{W}^{m,p}(\mathbb{T}^2)$ (resp. $\mathbb{W}^{m,p}(\mathbb{T}^2)$) and $\mathrm{H}^m(\mathbb{T}^2)$ (resp. $\mathbb{H}^m(\mathbb{T}^2)$) will be denoted by $\|\cdot\|_p$, $\|\cdot\|_{\mathrm{W}^{m,p}}$ (resp. $\|\cdot\|_{\mathbb{W}^{m,p}}$) and $\|\cdot\|_{\mathrm{H}^m}$ (resp. $\|\cdot\|_{\mathbb{H}^m}$), respectively. We set $\|\cdot\|_{\mathrm{W}^{0,p}} = \mathrm{L}^p(\mathbb{T}^2)$ (resp. $\|\cdot\|_{\mathbb{W}^{0,p}} = \mathbb{L}^p(\mathbb{T}^2)$).

Let $\dot{\mathrm{C}}^\infty(\mathbb{T}^2;\mathbb{R}^2)$ denote the space of all infinitely differentiable functions defined on $\mathbb{T}^2$ with values in $\mathbb{R}^2$ such that $\int_{\mathbb{T}^2} \boldsymbol{v}(x)\mathrm{d}x = \mathbf{0}$. We notice that the functions in

$\dot{\mathrm{C}}^\infty(\mathbb{T}^2;\mathbb{R}^2)$ satisfy the periodic boundary condition $\boldsymbol{v}(x+\mathrm{L}e_i)=\boldsymbol{v}(x)$. The Sobolev space $\dot{\mathbb{H}}^s(\mathbb{T}^2)$, $s\in\mathbb{N}_0=:\mathbb{N}\cup\{0\}$, is defined as the completion of $\dot{\mathrm{C}}^\infty(\mathbb{T}^2;\mathbb{R}^2)$ with respect to the following Sobolev norm

$$\|\boldsymbol{v}\|_{\mathbb{H}^s}:=\left(\sum_{0\leq|\alpha|\leq s}\|\mathrm{D}^\alpha\boldsymbol{v}\|^2_{\mathbb{L}^2(\mathbb{T}^2)}\right)^{1/2}.$$

According to [24, Proposition 5.39], we have

$$\dot{\mathbb{H}}^s(\mathbb{T}^2)$$
$$=\left\{\boldsymbol{v}:\boldsymbol{v}=\sum_{\boldsymbol{k}\in\mathbb{Z}^2}\boldsymbol{v}_{\boldsymbol{k}}e^{2\pi i\boldsymbol{k}\cdot x/\mathrm{L}},\boldsymbol{v}_0=\mathbf{0},\ \bar{\boldsymbol{v}}_{\boldsymbol{k}}=\boldsymbol{v}_{-\boldsymbol{k}},\ \|\boldsymbol{v}\|^2_{\dot{\mathbb{H}}^s}:=\sum_{\boldsymbol{k}\in\mathbb{Z}^2}|\boldsymbol{k}|^{2s}|\boldsymbol{v}_{\boldsymbol{k}}|^2<\infty\right\},$$

and from [24, Proposition 5.38], we infer that $\|\cdot\|_{\dot{\mathbb{H}}^s}$ defines a norm on the space $\|\cdot\|_{\dot{\mathbb{H}}^s}$, which is equivalent to the induced standard Sobolev norm $\|\cdot\|_{\mathbb{H}^s}$ on $\dot{\mathbb{H}}^s(\mathbb{T}^2)$. The zero mean condition provides the well-known *Poincaré inequality*,

$$\lambda_1\int_{\mathbb{T}^2}|\boldsymbol{v}(x)|^2\mathrm{d}x\leq\int_{\mathbb{T}^2}|\nabla\boldsymbol{v}(x)|^2\mathrm{d}x,\ \text{for any }\boldsymbol{v}\in\dot{\mathbb{H}}^1(\mathbb{T}^2),\tag{6}$$

where $\lambda_1=\frac{4\pi^2}{\mathrm{L}^2}$ ([24, Lemma 5.40]). Let us define $\mathcal{V}:=\{\boldsymbol{v}\in\dot{\mathrm{C}}^\infty(\mathbb{T}^2;\mathbb{R}^2):\nabla\cdot\boldsymbol{v}=0\}$. We denote by $\mathbb{H}$ the closure of $\mathcal{V}$ in the Lebesgue spaces $\dot{\mathbb{L}}^2(\mathbb{T}^2)$; and by $\mathbb{V}$ the closure of $\mathcal{V}$ in the Sobolev space $\dot{\mathbb{H}}^1(\mathbb{T}^2)$.

The spaces $\mathbb{H}$ and $\mathbb{V}$ are Hilbert spaces with inner products defined by

$$(\boldsymbol{u},\boldsymbol{v}):=\int_{\mathbb{T}^2}\boldsymbol{u}(x)\cdot\boldsymbol{v}(x)\mathrm{d}x=\sum_{i=1}^2\int_{\mathbb{T}^2}\boldsymbol{u}^i(x)\boldsymbol{v}^i(x)\,\mathrm{d}x,\qquad\text{for all }\boldsymbol{u},\boldsymbol{v}\in\mathbb{H},\tag{7}$$

$$(\boldsymbol{u},\boldsymbol{v})_{\mathbb{V}}:=(\nabla\boldsymbol{u},\nabla\boldsymbol{v}),\qquad\text{for all }\boldsymbol{u},\boldsymbol{v}\in\mathbb{V}.\tag{8}$$

We denote the corresponding norms by $\|\cdot\|_2$ and $\|\cdot\|_{\mathbb{V}}$, respectively.

Let us introduce the Banach space $(\mathbb{X}:=\mathbb{W}^{1,4}(\mathbb{T}^2)\cap\mathbb{V},\|\cdot\|_{\mathbb{X}})$ with $\|\cdot\|_{\mathbb{X}}:=\|\cdot\|_{\mathbb{W}^{1,4}}+\|\cdot\|_{\mathbb{V}}$. We recall that $\mathbb{W}^{1,4}(\mathbb{T}^2)$ endowed with the usual norm $\|\cdot\|_{\mathbb{W}^{1,4}}$

$$\|\boldsymbol{w}\|^4_{\mathbb{W}^{1,4}}=\int_{\mathbb{T}^2}|\boldsymbol{w}(x)|^4\mathrm{d}x+\int_{\mathbb{T}^2}|\nabla\boldsymbol{w}(x)|^4\mathrm{d}x$$

is a Banach space. We represent by $\langle\cdot,\cdot\rangle$ the duality relation between the spaces $\mathbb{V}$ and its dual $\mathbb{V}'$ as well as $\mathbb{X}$ and its dual $\mathbb{X}'$.

Next, let us introduce the scalar product between two matrices $A:B=\mathrm{Tr}(AB^T)$ and denote $|A|^2:=A:A$. The divergence of a matrix $A\in\mathcal{M}_{2\times2}(E)$ is given by $\left\{\mathrm{div}(A)_i\right\}_{i=1}^2=\left\{\sum_{j=1}^2\partial_ja_{ij}\right\}_{i=1}^2$. We recall that

$$(A,B)=\int_{\mathbb{T}^2}A(x):B(x)\,\mathrm{d}x;\qquad\text{for all }A,B\in\mathcal{M}_{2\times2}(\mathbb{L}^2(\mathbb{T}^2)).$$

Throughout the article, we denote by C generic constant, which may vary from line to line.

2.2. **Linear and nonlinear operators.** Let $\mathcal{P} : \dot{\mathbb{L}}^2(\mathbb{T}^2) \to \mathbb{H}$ be the Helmholtz-Hodge (or Leray) orthogonal projection (cf. [2]). We define the Stokes operator

$$\mathcal{A}\boldsymbol{u} := -\mathcal{P}\Delta\boldsymbol{u}, \ \boldsymbol{u} \in \mathrm{D}(\mathcal{A}),$$

where $\mathrm{D}(\mathcal{A}) = \big\{\boldsymbol{u} \in \dot{\mathbb{H}}^2(\mathbb{T}^2) : \nabla\cdot\boldsymbol{u} = 0\big\}$. It should be noted that $\mathcal{P}$ and Δ commutes in a torus, see [25, Lemma 2.9]. Therefore, in the sequel, for $\boldsymbol{u} \in \mathrm{D}(\mathcal{A})$, we will write

$$\Upsilon(\boldsymbol{u}) := \boldsymbol{u} + \alpha_1\mathcal{A}\boldsymbol{u} = \boldsymbol{u} - \alpha_1\Delta\boldsymbol{u}.$$

Given $\boldsymbol{u} \in \mathrm{D}(\mathcal{A})$ with Fourier expansion $\boldsymbol{u} = \sum\limits_{\boldsymbol{k}\in\mathbb{Z}^2} e^{2\pi i\boldsymbol{k}\cdot x/\mathrm{L}}\boldsymbol{u}_{\boldsymbol{k}}$, one obtains

$$-\Delta\boldsymbol{u} = \frac{4\pi^2}{\mathrm{L}^2}\sum_{\boldsymbol{k}\in\mathbb{Z}^2} e^{2\pi i\boldsymbol{k}\cdot x/\mathrm{L}}|\boldsymbol{k}|^2\boldsymbol{u}_{\boldsymbol{k}}.$$

Since $\mathcal{A}^{-1}$ is a compact self-adjoint operator in $\mathbb{H}$, there exists an orthonormal basis in $\mathbb{H}$ of eigenfunctions $\{\boldsymbol{w}_i\}_{i\in\mathbb{N}} \subset \dot{\mathrm{C}}^\infty(\mathbb{T}^2;\mathbb{R}^2)$ of the Stokes operator; namely we have $\mathcal{A}\boldsymbol{w}_i = \lambda_i\boldsymbol{w}_i$, for $i = 1, 2, \ldots$, where the eigenvalues of $\mathcal{A}$ verify $0 < \lambda_1 \le \lambda_2 \le \ldots \to \infty$. Note that $\lambda_1 = \frac{4\pi^2}{\mathrm{L}^2}$ is the smallest eigenvalue of $\mathcal{A}$ appearing in the Poincaré inequality (6). Moreover, the following relation holds

$$\lambda_i(\boldsymbol{w}_i, \boldsymbol{v}) = (\mathcal{A}\boldsymbol{w}_i, \boldsymbol{v}) = (-\Delta\boldsymbol{w}_i, \boldsymbol{v}) = (\nabla\boldsymbol{w}_i, \nabla\boldsymbol{v}) = (\boldsymbol{w}_i, \boldsymbol{v})_{\mathbb{V}}, \quad \text{for all } \boldsymbol{v} \in \mathbb{V}. \tag{9}$$

Hence $\{\boldsymbol{w}_i\}_{i\in\mathbb{N}}$ is an orthogonal basis in $\mathbb{V}$.

We recall that the operators $\mathcal{A}^\lambda$, $\lambda \in \mathbb{R}$, are well defined and

$$\mathrm{D}(\mathcal{A}^{s/2}) = \big\{\boldsymbol{u} \in \dot{\mathbb{H}}^s(\mathbb{T}^2) : \nabla\cdot\boldsymbol{u} = 0\big\}.$$

In addition, it can be shown that there exists a positive constant C such that $\|\mathcal{A}^{s/2}\boldsymbol{u}\|_2 = C\|\boldsymbol{u}\|_{\dot{\mathbb{H}}^s}$, for all $\boldsymbol{u} \in \mathrm{D}(\mathcal{A}^{s/2})$, $s \ge 0$ (see [24, Chapter 6]). Note that the operator $\mathcal{A}$ is a non-negative self-adjoint operator in $\mathbb{H}$ with a compact resolvent and

$$|(\mathcal{A}\boldsymbol{u}, \boldsymbol{v})| \le C\|\boldsymbol{u}\|_{\mathbb{V}}\|\boldsymbol{v}\|_{\mathbb{V}}, \ \text{ for all } \ \boldsymbol{u} \in \mathrm{D}(\mathcal{A}), \ \boldsymbol{v} \in \mathbb{V}, \ \text{then} \ \|\mathcal{A}\boldsymbol{u}\|_{\mathbb{V}'} \le \|\boldsymbol{u}\|_{\mathbb{V}}. \tag{10}$$

Therefore, there exists a unique extension of $\mathcal{A}$ (denoted by the same symbol) verifying

$$\mathcal{A} : \mathbb{V} \to \mathbb{V}', \qquad \langle\mathcal{A}\boldsymbol{u}, \boldsymbol{v}\rangle := (\nabla\boldsymbol{u}, \nabla\boldsymbol{v}), \qquad \forall\boldsymbol{u}, \ \boldsymbol{v} \in \mathbb{V}.$$

Next, we define the trilinear form $b(\cdot,\cdot,\cdot) : \mathbb{V}\times\mathbb{V}\times\mathbb{V} \to \mathbb{R}$ by

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \int_{\mathbb{T}^2}(\boldsymbol{u}(x)\cdot\nabla)\boldsymbol{v}(x)\cdot\boldsymbol{w}(x)\mathrm{d}x = \sum_{i,j=1}^{2}\int_{\mathbb{T}^2}\boldsymbol{u}^i(x)\frac{\partial\boldsymbol{v}^j(x)}{\partial x_i}\boldsymbol{w}^j(x)\mathrm{d}x.$$

If $\boldsymbol{u}, \boldsymbol{v}$ are such that the linear map $b(\boldsymbol{u}, \boldsymbol{v}, \cdot)$ is continuous on $\mathbb{V}$, the corresponding element of $\mathbb{V}'$ is denoted by $\mathrm{B}(\boldsymbol{u}, \boldsymbol{v})$. We represent $\mathrm{B}(\boldsymbol{v}) = \mathrm{B}(\boldsymbol{v}, \boldsymbol{v}) = \mathcal{P}[(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}]$. Using an integration by parts, it is immediate that

$$\begin{cases} b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = 0, & \text{for all } \ \boldsymbol{u}, \boldsymbol{v} \in \mathbb{V}, \\ b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = -b(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}), & \text{for all } \ \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{V}. \end{cases} \tag{11}$$

For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{V}$, we have

$$\begin{aligned} |\langle\mathrm{B}(\boldsymbol{u}) - \mathrm{B}(\boldsymbol{v}), \boldsymbol{w}\rangle| &\le |b(\boldsymbol{u}, \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{w})| + |b(\boldsymbol{u}-\boldsymbol{v}, \boldsymbol{v}, \boldsymbol{w})| \\ &\le \big[\|\boldsymbol{u}\|_4\|\nabla(\boldsymbol{u}-\boldsymbol{v})\|_2 + \|\boldsymbol{u}-\boldsymbol{v}\|_4\|\nabla\boldsymbol{v}\|_2\big]\|\boldsymbol{w}\|_4, \end{aligned} \tag{12}$$

for all $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\mathbb{V}$. Thus the operator $\mathrm{B}(\cdot):\mathbb{V}\to\mathbb{V}'$ is locally Lipschitz. In addition, the operator B enjoys the following important orthogonality property (see [32, Lemma 3.1, p. 404]):

$$(\mathrm{B}(\boldsymbol{v},\boldsymbol{v}),\mathcal{A}\boldsymbol{v})=b(\boldsymbol{v},\boldsymbol{v},\mathcal{A}\boldsymbol{v})=0,\ \text{for any } \boldsymbol{v}\in\mathrm{D}(\mathcal{A}). \tag{13}$$

Let us now define the operator $\mathcal{J}(\boldsymbol{v}):=-\mathcal{P}[\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\boldsymbol{v}))]$. Note that for $\boldsymbol{v}\in\mathbb{X}$, we have

$$\|\mathcal{J}(\boldsymbol{v})\|_{\mathbb{X}'}\leq C\|\boldsymbol{v}\|_{\mathbb{X}}^2,$$

and hence the map $\mathcal{J}(\cdot):\mathbb{X}\to\mathbb{X}'$. For $\boldsymbol{u},\boldsymbol{v}\in\mathbb{X}$, we have

$$\begin{aligned}|\langle\mathcal{J}(\boldsymbol{u})-\mathcal{J}(\boldsymbol{v}),\boldsymbol{w}\rangle|&=\left|\frac{1}{2}\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{u}-\boldsymbol{v})\mathrm{A}(\boldsymbol{u})+\mathrm{A}(\boldsymbol{v})\mathrm{A}(\boldsymbol{u}-\boldsymbol{v})]:\mathrm{A}(\boldsymbol{w})\mathrm{d}x\right|\\&\leq C\big[\|\mathrm{A}(\boldsymbol{u})\|_4+\|\mathrm{A}(\boldsymbol{v})\|_4\big]\|\mathrm{A}(\boldsymbol{u}-\boldsymbol{v})\|_4\|\mathrm{A}(\boldsymbol{w})\|_2,\end{aligned} \tag{14}$$

for all $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\mathbb{X}$. Thus the operator $\mathcal{J}(\cdot):\mathbb{X}\to\mathbb{X}'$ is locally Lipschitz.

Finally, we define the operator $\mathcal{K}(\boldsymbol{v}):=-\mathcal{P}[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\boldsymbol{v}))]$. It is immediate that

$$\langle\mathcal{K}(\boldsymbol{v}),\boldsymbol{v}\rangle=\frac{1}{2}\|\mathrm{A}(\boldsymbol{v})\|_4^4.$$

Note that for $\boldsymbol{v}\in\mathbb{X}$, we have

$$\|\mathcal{K}(\boldsymbol{v})\|_{\mathbb{X}'}\leq C\|\boldsymbol{v}\|_{\mathbb{X}}^3,$$

and hence the map $\mathcal{K}(\cdot):\mathbb{X}\to\mathbb{X}'$. For $\boldsymbol{u},\boldsymbol{v}\in\mathbb{X}$, we infer

$$\begin{aligned}|\langle\mathcal{K}(\boldsymbol{u})-\mathcal{K}(\boldsymbol{v}),\boldsymbol{w}\rangle|&=\left|\frac{1}{2}\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{u}-\boldsymbol{v}):\mathrm{A}(\boldsymbol{u})+\mathrm{A}(\boldsymbol{v}):\mathrm{A}(\boldsymbol{u}-\boldsymbol{v})]\mathrm{A}(\boldsymbol{u}):\mathrm{A}(\boldsymbol{w})\mathrm{d}x\right.\\&\qquad\left.+\frac{1}{2}\int_{\mathbb{T}^2}|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\boldsymbol{u}-\boldsymbol{v}):\mathrm{A}(\boldsymbol{w})\mathrm{d}x\right|\\&\leq C\big[\|\mathrm{A}(\boldsymbol{u})\|_4^2+\|\mathrm{A}(\boldsymbol{v})\|_4^2\big]\|\mathrm{A}(\boldsymbol{u}-\boldsymbol{v})\|_4\|\mathrm{A}(\boldsymbol{w})\|_4,\end{aligned} \tag{15}$$

for all $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\mathbb{X}$. Thus the operator $\mathcal{K}(\cdot):\mathbb{X}\to\mathbb{X}'$ is locally Lipschitz.

2.3. **Linear stochastic equation.** In order to establish global-in-time solution of system (3), we decompose system (3) in a linear equation which involves the noise and a random partial differential equation. Namely, we write $\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{z}$ such that $\boldsymbol{z}$ solves the following stochastic linear system:

$$\begin{cases}\mathrm{d}(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}-\theta\Delta(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}\mathrm{d}t+(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}\mathrm{d}t=\mathrm{dW},\\ \boldsymbol{z}(0)=\boldsymbol{0},\end{cases} \tag{16}$$

where $\theta>0$ is a given constant, W is $\mathbb{R}^2$-valued two-sided trace-class GG^*-Wiener process with spatial mean zero and divergence-free, and $\boldsymbol{u}$ solves the following nonlinear random system:

$$\begin{cases} \partial_t(\boldsymbol{u}-\alpha_1\Delta\boldsymbol{u})-\nu\Delta\boldsymbol{u}+((\boldsymbol{u}+\boldsymbol{z})\cdot\nabla)(\boldsymbol{u}+\boldsymbol{z}-\alpha_1\Delta\boldsymbol{u}-\alpha_1\Delta\boldsymbol{z}) \\ +\sum_{j=1}^{2}[\boldsymbol{u}+\boldsymbol{z}-\alpha_1\Delta\boldsymbol{u}-\alpha_1\Delta\boldsymbol{z}]^j\nabla[\boldsymbol{u}+\boldsymbol{z}]^j \\ -(\alpha_1+\alpha_2)\mathrm{div}[\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})^2]-\beta\mathrm{div}[|\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})|^2\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})] \\ =-\nabla p+\boldsymbol{f}+\boldsymbol{z}+(\nu-\theta-\alpha_1)\Delta\boldsymbol{z}+\theta\alpha_1\Delta^2\boldsymbol{z}, & \text{in } \mathbb{T}^2\times(0,\infty), \\ \mathrm{div}\ \boldsymbol{u}=0, & \text{in } \mathbb{T}^2\times[0,\infty), \\ \boldsymbol{u}(x,0)=\boldsymbol{v}_0(x), & \text{in } \mathbb{T}^2, \end{cases} \tag{17}$$

with periodic boundary conditions.

Remark 2.1. Here, $\boldsymbol{z}$ is divergence-free due to the assumptions on the noise W, so there is no need to explicitly include the condition div $\boldsymbol{z}=0$ in (16). Therefore, we denote by p the pressure term associated with $\boldsymbol{u}$.

In view of the factorization method, one can obtain the regularity of $\boldsymbol{z}$ on a given stochastic basis $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P})$ with $\{\mathcal{F}_t\}_{t\in[0,T]}$ being the usual filtration. Particularly, we prove the following result for regularity of $\boldsymbol{z}$ using the similar arguments as in [13, Theorem 5.16].

Proposition 2.2. *Suppose that* $\mathrm{Tr}\left((-\Delta)^{3-2\gamma}GG^*\right)<\infty$ *for* $\gamma\in(0,\frac{1}{2})$. *Then for* $p\geq 2$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|\boldsymbol{z}(t)\|^p_{\dot{\mathbb{H}}^5}\right]\leq(p-1)^{\frac{p}{2}}L^p, \tag{18}$$

where $L\geq 1$ *depends on* $\mathrm{Tr}\left((-\Delta)^{3-2\gamma}GG^*\right)$, γ *and* θ, *and is independent of* p. *In addition, for any given* $\hat{c}>0$, *we can choose*

$$\theta>\hat{c}+\frac{\|GG^*\|_{Op}}{\lambda_1},$$

where $\|GG^*\|_{Op}$ *is the operator norm of operator* GG^* *from* $\mathbb{H}$ *to* $\mathbb{H}$ *and* λ_1 *is the smallest eigenvalue of Stokes operator, such that the following exponential estimate hold:*

$$\mathbb{E}\left[\exp\left\{2\hat{c}\int_0^t\|\boldsymbol{z}(s)\|^2_{\dot{\mathbb{H}}^3}\mathrm{d}s\right\}\right]<+\infty, \tag{19}$$

for all $t\in[0,T]$.

Remark 2.3. Note that $\dot{\mathbb{H}}^5$-regularity of $\boldsymbol{z}$ in (18) is required to obtain the $\dot{\mathbb{H}}^3$-regularity of $\boldsymbol{u}$ (see Subsection 3.1.6 below).

Proof of Proposition 2.2. The proof is divided into two steps.

Step I. *Proof of* (18). Let us recall that the unique stationary solution to (16) has the explicit form

$$\Upsilon(\boldsymbol{z}(t))=\int_0^t e^{-(t-r)}S(t-r)\mathrm{d}\mathrm{W}(r),$$

where $S(t)=e^{\theta\Delta t}$ denotes an analytic semigroup generated by the operator $\theta\Delta$. By [17, Section 2.1], the Wiener process W can be written as $\mathrm{W}=\sum_{k\in\mathbb{N}}\sqrt{c_k}\beta_k\boldsymbol{e}_k$ for an orthogonal basis $\{\boldsymbol{e}_k\}_{k\in\mathbb{N}}$ of $\mathbb{H}$ consisting of eigenfunctions of GG^* with corresponding eigenvalues c_k, and the coefficient satisfy $\sum\limits_{k\in\mathbb{N}} c_k < \infty$, where $\{\beta_k\}_{k\in\mathbb{N}}$ is a sequence of mutually independent standard real-valued Brownian motions. Also, in view of our assumption, we have (see [13, Appendix C])

$$\mathrm{Tr}\left((-\Delta)^{3-2\gamma}GG^*\right) = \sum_{k=1}^{\infty} c_k \left\|(-\Delta)^{\frac{3}{2}-\gamma}\boldsymbol{e}_k\right\|_2^2 < \infty, \quad \text{for } \ \gamma \in \left(0, \frac{1}{2}\right). \tag{20}$$

Then it holds for $\gamma \in \left(0, \frac{1}{2}\right)$ and $t \geq s$,

$$\begin{aligned}
&\mathbb{E}\left[\|\Upsilon(\boldsymbol{z}(t)) - \Upsilon(\boldsymbol{z}(s))\|_2^2\right] \\
&= \mathbb{E}\left[\left\|\int\limits_s^t e^{-(t-r)}S(t-r)\mathrm{dW}(r)\right\|_2^2\right] \\
&\quad + \mathbb{E}\left[\left\|\int\limits_0^s \left[e^{-(t-r)}S(t-r) - e^{-(s-r)}S(s-r)\right]\mathrm{dW}(r)\right\|_2^2\right] \\
&= \sum_{k=1}^{\infty} c_k \int\limits_s^t e^{-2(t-r)}\left\|S(t-r)\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r \\
&\quad + \sum_{k=1}^{\infty} c_k \int\limits_0^s \left\|\left[e^{-(t-r)}S(t-r) - e^{-(s-r)}S(s-r)\right]\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r \\
&\leq \sum_{k=1}^{\infty} c_k \int\limits_s^t e^{-2(t-r)}\left\|S(t-r)\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r \\
&\quad + \sum_{k=1}^{\infty} c_k \int\limits_0^s \left[e^{-(t-r)} - e^{-(s-r)}\right]^2 \left\|S(t-r)\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r \\
&\quad + \sum_{k=1}^{\infty} c_k \int\limits_0^s e^{-2(s-r)}\left\|\left[S(t-r) - S(s-r)\right]\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r \\
&\leq C\Bigg[\mathrm{Tr}(GG^*)\left\{\int\limits_s^t e^{-2(t-r)}\mathrm{d}r + \int\limits_0^s \left[e^{-(t-r)} - e^{-(s-r)}\right]^2 \mathrm{d}r\right\} \\
&\quad + \sum_{k=1}^{\infty} c_k \int\limits_0^s e^{-2(s-r)}\left\|\left[S(t-r) - S(s-r)\right]\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r\Bigg] \\
&\leq C\left[\mathrm{Tr}(GG^*)(t-s)^{2\gamma} + \sum_{k=1}^{\infty} c_k \int\limits_0^s e^{-2(s-r)}\left\|\left(\int_{s-r}^{t-r}(-\theta\Delta)S(\tau)\mathrm{d}\tau\right)\boldsymbol{e}_k\right\|_2^2 \mathrm{d}r\right] \\
&\leq C\left[\mathrm{Tr}(GG^*)(t-s)^{2\gamma} + \mathrm{Tr}(GG^*)\int\limits_0^s e^{-2(s-r)}\left|\int_{s-r}^{t-r}\frac{1}{\tau}\mathrm{d}\tau\right|^2 \mathrm{d}r\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq C\left[(t-s)^{2\gamma}+\int_0^s e^{-2(s-r)}(s-r)^{-2\gamma}\left|\int_{s-r}^{t-r}\tau^{\gamma-1}\mathrm{d}\tau\right|^2\mathrm{d}r\right]\\
&= C\left[(t-s)^{2\gamma}+\frac{1}{\gamma^2}\int_0^s e^{-2(s-r)}(s-r)^{-2\gamma}\left[(t-r)^{\gamma}-(s-r)^{\gamma}\right]^2\mathrm{d}r\right]\\
&\leq C(t-s)^{2\gamma}\left[1+\int_0^s e^{-2(s-r)}(s-r)^{-2\gamma}\mathrm{d}r\right]\\
&\leq C(t-s)^{2\gamma},
\end{aligned}
$$

where the constant C depends only on the semigroup $S(\cdot)$, $\mathrm{Tr}(GG^*)$, θ and γ but it is independent of time. Again, it holds for $\gamma\in(0,\frac{1}{2})$, $\xi\in(\gamma,\frac{1}{2})$ and $t\geq s$,

$$
\begin{aligned}
&\mathbb{E}\left[\left\|(-\Delta)^{\frac{3}{2}}\Upsilon(\boldsymbol{z}(t))-(-\Delta)^{\frac{3}{2}}\Upsilon(\boldsymbol{z}(s))\right\|_2^2\right]\\
&=\mathbb{E}\left[\left\|\int_s^t e^{-(t-r)}(-\Delta)^{\frac{3}{2}}S(t-r)\mathrm{dW}(r)\right\|_2^2\right]\\
&\quad+\mathbb{E}\left[\left\|\int_0^s\left[e^{-(t-r)}(-\Delta)^{\frac{3}{2}}S(t-r)-e^{-(s-r)}(-\Delta)^{\frac{3}{2}}S(s-r)\right]\mathrm{dW}(r)\right\|_2^2\right]\\
&=\sum_{k=1}^{\infty}c_k\int_s^t e^{-2(t-r)}\left\|(-\Delta)^{\frac{3}{2}}S(t-r)\boldsymbol{e}_k\right\|_2^2\mathrm{d}r\\
&\quad+\sum_{k=1}^{\infty}c_k\int_0^s\left\|\left[e^{-(t-r)}(-\Delta)^{\frac{3}{2}}S(t-r)-e^{-(s-r)}(-\Delta)^{\frac{3}{2}}S(s-r)\right]\boldsymbol{e}_k\right\|_2^2\mathrm{d}r\\
&\leq C\Bigg[\sum_{k=1}^{\infty}c_k\int_s^t e^{-2(t-r)}\left\|(-\Delta)^{\frac{3}{2}-\gamma}(-\Delta)^{\gamma}S(t-r)\boldsymbol{e}_k\right\|_2^2\mathrm{d}r\\
&\quad+\sum_{k=1}^{\infty}c_k\int_0^s e^{-2(t-r)}\left\|(-\Delta)^{\frac{3}{2}-\gamma}(-\Delta)^{\gamma}\left[S(t-r)-S(s-r)\right]\boldsymbol{e}_k\right\|_2^2\mathrm{d}r\\
&\quad+\sum_{k=1}^{\infty}c_k\int_0^s\left[e^{-(t-r)}-e^{-(s-r)}\right]^2\left\|(-\Delta)^{\frac{3}{2}-\gamma}(-\Delta)^{\gamma}S(s-r)\boldsymbol{e}_k\right\|_2^2\mathrm{d}r\Bigg]\\
&\leq C\Bigg[\mathrm{Tr}((-\Delta)^{3-2\gamma}GG^*)\int_s^t e^{-2(t-r)}(t-r)^{-2\gamma}\mathrm{d}r\\
&\quad+\mathrm{Tr}((-\Delta)^{3-2\gamma}GG^*)\int_0^s\left[e^{-(t-r)}-e^{-(s-r)}\right]^2(s-r)^{-2\gamma}\mathrm{d}r
\end{aligned}
$$

$$+\sum_{k=1}^{\infty} c_k \int\limits_0^s e^{-2(t-r)} \left\| (-\Delta)^{\frac{3}{2}-\gamma}(-\Delta)^{\gamma} \left(\int_{s-r}^{t-r} \frac{\mathrm{d}}{\mathrm{d}\tau} S(\tau)\mathrm{d}\tau \right) \boldsymbol{e}_k \right\|_2^2 \mathrm{d}r \Bigg]$$

$$\leq C\Bigg[\sum_{k=1}^{\infty} c_k \int\limits_0^s e^{-2(t-r)} \left\| (-\Delta)^{\frac{3}{2}-\gamma}(-\theta\Delta)^{\gamma} \left(\int_{s-r}^{t-r} (\theta\Delta) S(\tau)\mathrm{d}\tau \right) \boldsymbol{e}_k \right\|_2^2 \mathrm{d}r$$

$$+ (t-s)^{2(\xi-\gamma)} \Bigg]$$

$$\leq C\left[(t-s)^{2(\xi-\gamma)} + \int\limits_0^s e^{-2(t-r)} \left| \int_{s-r}^{t-r} \tau^{-1-\gamma}\mathrm{d}\tau \right|^2 \mathrm{d}r \right]$$

$$\leq C\left[(t-s)^{2(\xi-\gamma)} + \int\limits_0^s e^{-2(t-r)} (s-r)^{-2\xi} \left| \int_{s-r}^{t-r} \tau^{\xi-1-\gamma}\mathrm{d}\tau \right|^2 \mathrm{d}r \right]$$

$$\leq C\left[(t-s)^{2(\xi-\gamma)} + \int\limits_0^s e^{-2(t-r)} (s-r)^{-2\xi} \left[(t-r)^{\xi-\gamma} - (s-r)^{\xi-\gamma} \right]^2 \mathrm{d}r \right]$$

$$\leq C\left[(t-s)^{2(\xi-\gamma)} + (t-s)^{2(\xi-\gamma)} \int\limits_0^s e^{-2(t-r)} (s-r)^{-2\xi} \mathrm{d}r \right]$$

$$\leq C(t-s)^{2(\xi-\gamma)},$$

where the constant $C > 0$ depends only on the semigroup $S(\cdot)$, $\mathrm{Tr}((-\Delta)^{3-2\gamma}GG^*)$, θ, γ and ξ but it is independent of time, and we have also used the properties of semigroup $S(\cdot)$ from [21, Theorems 5.2 and 6.13]. Using Gaussianity (see for e.g. [19, Subsection 5.4]), we have

$$\mathbb{E}\left[\|\boldsymbol{z}(t) - \boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^5}^p \right] \leq (p-1)^{\frac{p}{2}} \left(\mathbb{E}\left[\|\boldsymbol{z}(t) - \boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^5}^2 \right] \right)^{\frac{p}{2}}.$$

By Kolmogorov's continuity criterion (see [13, Theorem 3.3]), we obtain (18) as desired.

Step II. *Proof of* (19). Applying Itô formula to the process $\|(\mathrm{I} - \alpha_1\Delta)\boldsymbol{z}(\cdot)\|_2^2$ and integrating form 0 to t, we obtain

$$\|(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(t)\|_2^2 + 2\theta \int_0^t \|(-\Delta)^{\frac{1}{2}}(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2 \mathrm{d}s + 2\int_0^t \|(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2 \mathrm{d}s$$

$$= 2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s), \mathrm{dW}(s)) + \mathrm{Tr}(GG^*)t$$

$$= 2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s), \mathrm{dW}(s))\mathrm{d}s - 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2 \mathrm{d}s + \mathrm{Tr}(GG^*)t$$

$$+ 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2 \mathrm{d}s$$

$$= 2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s), \mathrm{dW}(s))\mathrm{d}s - 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2 \mathrm{d}s + \mathrm{Tr}(GG^*)t$$

$$+ 2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s), GG^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s))\mathrm{d}s$$

$$
\begin{aligned}
&\leq 2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s),\mathrm{dW}(s))\mathrm{d}s - 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\mathrm{d}s + \mathrm{Tr}(GG^*)t \\
&\quad + 2\|GG^*\|_{Op}\int_0^t \|(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\mathrm{d}s \\
&\leq 2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s),\mathrm{dW}(s))\mathrm{d}s - 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\mathrm{d}s + \mathrm{Tr}(GG^*)t \\
&\quad + \frac{2\|GG^*\|_{Op}}{\lambda_1}\int_0^t \|(-\Delta)^{\frac12}(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\mathrm{d}s, \qquad (21)
\end{aligned}
$$

where $\|GG^*\|_{Op}$ denotes the operator norm of operator GG^* from $\mathbb{H}$ to $\mathbb{H}$. This gives

$$
\begin{aligned}
&\exp\left\{2\left(\theta-\frac{\|GG^*\|_{Op}}{\lambda_1}\right)\int_0^t \|(-\Delta)^{\frac12}(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\mathrm{d}s\right\} \\
&\leq \exp\left\{2\int_0^t ((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s),\mathrm{dW}(s)) - 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\right\} e^{\mathrm{Tr}(GG^*)t}. \qquad (22)
\end{aligned}
$$

Since $\mathcal{E}(t) := \exp\left\{2\int_0^t((\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s),\mathrm{dW}(s)) - 2\int_0^t \|G^*(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\right\}$ is a non-negative local martingale, it follows by Fatou's Lemma that $\mathcal{E}$ is a supermartingale (see [17, pp. 253]), and since $\mathcal{E}(0)=1$, we have

$$
\mathbb{E}\left[\mathcal{E}(t)\right] \leq \mathbb{E}\left[\mathcal{E}(0)\right] = 1. \qquad (23)
$$

Hence by (22) and (23), we have

$$
\mathbb{E}\left[\exp\left\{2\left(\theta-\frac{\|GG^*\|_{Op}}{\lambda_1}\right)\int_0^t \|(-\Delta)^{\frac12}(\mathrm{I}-\alpha_1\Delta)\boldsymbol{z}(s)\|_2^2\mathrm{d}s\right\}\right] \leq e^{\mathrm{Tr}(GG^*)t}. \qquad (24)
$$

Finally, for any given $\hat{c}>0$, one can choose θ sufficiently large satisfying

$$
\hat{c} < \theta - \frac{\|GG^*\|_{Op}}{\lambda_1},
$$

and obtain

$$
\mathbb{E}\left[\exp\left\{2\hat{c}\int_0^t \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\right\}\right] \leq e^{\mathrm{Tr}(GG^*)t} < +\infty. \qquad (25)
$$

This completes the proof. □

2.4. **Control problem.** Our main goal is to control the solution of the system (3) by a stochastic distributed force $\boldsymbol{f}$. The control variable $\boldsymbol{f}$ belongs to the set $\mathcal{F}_{ad} \subset \mathcal{X} := \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))) \cap \mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>4$) of admissible controls, which is given by

$$
\mathcal{F}_{ad} := \{\boldsymbol{f}\in\mathcal{X} : \|\boldsymbol{f}(\omega)\|_{\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))} \leq K\varrho(\omega), \text{ for a.s. } \omega\in\Omega\},
$$

where $0<K<+\infty$ and the random variable $\varrho\in\mathrm{L}^2(\Omega)$. Note that $\mathcal{F}_{ab}$ is closed and bounded in $\mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))$, and convex in $\mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))) \cap \mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (see Remark 2.4 below for more details).

Let $\boldsymbol{v}$ be the solution to system (3) corresponding to control $\boldsymbol{f}\in\mathcal{F}_{ad}$. We consider the cost functional $\mathrm{J}:\mathcal{F}_{ad}\to\mathbb{R}^+$ given by

$$
\mathrm{J}(\boldsymbol{f},\boldsymbol{v}) = \frac12\mathbb{E}\bigg[\int_0^T \|\boldsymbol{v}(t)-\boldsymbol{v}_d(t)\|_2^2\mathrm{d}t\bigg] + \frac{\lambda}{2}\mathbb{E}\bigg[\int_0^T \|\boldsymbol{f}(t)\|_2^2\mathrm{d}t\bigg], \qquad (26)
$$

where $\boldsymbol{v}_d \in \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ corresponds to a desired target field and any $\lambda > 0$. The control problem reads

$$\min_{\boldsymbol{f}\in\mathcal{F}_{ad}} \left\{ \mathrm{J}(\boldsymbol{f},\boldsymbol{v}) \ : \ \boldsymbol{v} \text{ is the solution of (3) with force } \boldsymbol{f} \right\}. \tag{27}$$

Remark 2.4 (Role and justification of the assumptions on the admissible control set)**.** In this remark, we explain the role of each assumption imposed on the admissible control set $\mathcal{F}_{ad}$ throughout the paper.

1. To derive the first-order necessary optimality condition satisfied by an optimal pair, it is essential to establish a duality relation between the solutions of the linearized system (129) and the corresponding adjoint system (191). The derivation of this duality relation requires suitable moment estimates for the solutions of both systems.

2. The well-posedness of the linearized system (129) and the adjoint system (191), together with the required moment bounds, relies on the regularity property (5) of the solution to the state equation (3). In turn, this regularity is obtained under the assumption

$$\mathcal{F}_{ad} \subset \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))).$$

3. Furthermore, the existence and uniqueness of solutions to the linearized and adjoint systems satisfying the desired moment estimates require

$$\mathcal{F}_{ad} \subset \mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))), \qquad p>2.$$

4. To prove that the solution of the linearized system (129) coincides with the Gâteaux derivative of the control-to-state mapping, we additionally require

$$\mathcal{F}_{ad} \subset \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))) \cap \mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))), \qquad p>4,$$

see the proof of Proposition 5.2.

5. The assumption that the $\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))$-norms of the elements of $\mathcal{F}_{\mathrm{ad}}$ are uniformly bounded by a random variable is required in the proof of the existence of an optimal pair; see Subsection 7.2.

6. The convexity of $\mathcal{F}_{ad}$ is used in the computation of the Gâteaux derivative of the cost functional J, which subsequently yields the first-order necessary optimality condition through admissible convex perturbations of the control, see Subsection 7.3.

7. Finally, the optimality condition is derived by considering admissible convex perturbations of an optimal control $\widetilde{\boldsymbol{f}} \in \mathcal{F}_{ad}$. Since the existence theorem (see Subsection 7.2 below) establishes that the optimal pair $(\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}})$ satisfies $\widetilde{\boldsymbol{f}} \in \mathcal{F}_{ad}$, the variational inequality obtained in Subsection 7.3 is valid over the admissible set. Thus, the resulting optimality condition characterizes an admissible optimal control.

Remark 2.5. In this article, the Lagrangian $\mathfrak{L}$ is given by

$$\mathfrak{L}(\cdot,\boldsymbol{f},\boldsymbol{v}) = \frac{1}{2}\|\boldsymbol{v}-\boldsymbol{v}_d\|_2^2 + \frac{\lambda}{2}\|\boldsymbol{f}\|_2^2.$$

Therefore, we have

$$\mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{v}}\mathfrak{L}(t,\boldsymbol{f}(t),\boldsymbol{v}(t)),\boldsymbol{y}(t))\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T (\boldsymbol{y}(t),\boldsymbol{v}(t)-\boldsymbol{v}_d(t))\mathrm{d}t\right],$$

$$\mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{f}}\mathfrak{L}(t,\boldsymbol{f}(t),\boldsymbol{v}(t)),\boldsymbol{g}(t))\mathrm{d}t\right] = \lambda\mathbb{E}\left[\int_0^T (\boldsymbol{g}(t),\boldsymbol{f}(t))\mathrm{d}t\right],$$

for any $\boldsymbol{y} \in \mathrm{L}^2(\Omega;\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}})))$ and $\boldsymbol{g} \in \mathcal{F}_{ad}$.

2.5. **Main results.** Our main result demonstrates the existence of a solution to the control problem and formulates the first-order necessary optimality conditions.

Theorem 2.6. *Assume that* $\boldsymbol{v}_0 \in \mathrm{D}(\mathcal{A}^{\frac{3}{2}})$. *Then the control problem* (27) *admits, at least, one optimal solution*

$$(\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}}) \in \mathcal{F}_{ad}\times \mathrm{L}^2(\Omega;\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))), \tag{28}$$

where $\widetilde{\boldsymbol{v}}$ *is the unique solution of* (3) *with* $\boldsymbol{f} = \widetilde{\boldsymbol{f}}$. *Moreover, there exists a unique solution* $\mathfrak{p}$ *of* (191) *with* $\boldsymbol{g} = \nabla_{\boldsymbol{v}}\mathfrak{L}(\cdot,\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}})$, *such that if* $\widetilde{\mathfrak{m}}$ *is the solution of* (129) *for* $\boldsymbol{v} = \widetilde{\boldsymbol{v}}$ *and* $\boldsymbol{\psi} = \boldsymbol{\psi} - \widetilde{\boldsymbol{f}}$, *the following duality property*

$$\mathbb{E}\left[\int_0^T (\boldsymbol{\psi}(t)-\widetilde{\boldsymbol{f}}(t),\widetilde{\mathfrak{p}}(t))\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{v}}\mathfrak{L}(t,\widetilde{\boldsymbol{f}}(t),\widetilde{\boldsymbol{v}}(t)),\widetilde{\mathfrak{m}}(t))\mathrm{d}t\right], \tag{29}$$

and the following optimality condition hold

$$\mathbb{E}\left[\int_0^T (\boldsymbol{\psi}(t)-\widetilde{\boldsymbol{f}}(t),\widetilde{\mathfrak{p}}(t)+\nabla_{\boldsymbol{f}}\mathfrak{L}(t,\widetilde{\boldsymbol{f}}(t),\widetilde{\boldsymbol{v}}(t)))\mathrm{d}t\right] \geq 0. \tag{30}$$

3. **Existence and uniqueness of global solutions of state equation** (3)**.** In this section, we establish the existence of unique global solution to system (3). Taking projection $\mathcal{P}$ to system (3), we obtain

$$\begin{cases} \mathrm{d}\Upsilon(\boldsymbol{v}) + \left\{\nu\mathcal{A}\boldsymbol{v} + \mathcal{P}[(\boldsymbol{v}\cdot\nabla)\Upsilon(\boldsymbol{v})] + \mathcal{P}\left[\sum_{j=1}^2[\Upsilon(\boldsymbol{v})]^j\nabla\boldsymbol{v}^j\right] - (\alpha_1+\alpha_2)\mathcal{J}(\boldsymbol{v}) - \beta\mathcal{K}(\boldsymbol{v})\right\}\mathrm{d}t \\ = \mathcal{P}\boldsymbol{f}\mathrm{d}t + \mathrm{dW}, \quad \text{for a.e. } t\in(0,T), \\ \boldsymbol{v}(0) = \boldsymbol{v}_0, \end{cases} \tag{31}$$

in $\mathbb{H}$. Let us first provide the definition of a unique global strong solution in the probabilistic sense to the system (3).

Definition 3.1. A $\mathrm{D}(\mathcal{A})$-valued $\{\mathscr{F}_t\}_{t\in[0,T]}$-adapted stochastic process $\boldsymbol{v}(\cdot)$ is called a *strong solution* to the system (3) if the following conditions are satisfied:

(i) the process $\{\boldsymbol{v}(t)\}_{t\in[0,T]}$ has trajectories in $\mathrm{C}([0,T];\mathrm{D}(\mathcal{A}))\cap\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))$, $\mathbb{P}$-a.s.,

(ii) the following equality holds for every $t\in[0,T]$ and for any $\psi\in\mathbb{H}$:

$$\begin{aligned}(\Upsilon(\boldsymbol{v}(t)),\psi) &= (\Upsilon(\boldsymbol{v}_0),\psi) - \int_0^t \langle \nu\mathcal{A}\boldsymbol{v}(s) - (\alpha_1+\alpha_2)\mathcal{J}(\boldsymbol{v}(s)) + \beta\mathcal{K}(\boldsymbol{v}(s)),\psi\rangle\mathrm{d}s \\ &\quad + \int_0^t b(\boldsymbol{v}(s),\psi(s),\Upsilon(\boldsymbol{v}(s)))\mathrm{d}s - \int_0^t b(\psi(s),\boldsymbol{v}(s),\Upsilon(\boldsymbol{v}(s)))\mathrm{d}s \\ &\quad + \int_0^t (\boldsymbol{f}(s),\psi)\mathrm{d}s + (\mathrm{W}(t),\psi), \ \mathbb{P}\text{-a.s.}\end{aligned} \tag{32}$$

Theorem 3.2. *Suppose that* $\mathrm{Tr}\left((-\Delta)^{3-2\gamma}GG^*\right) < \infty$ *for* $\gamma\in(0,\frac{1}{2})$, $\boldsymbol{v}_0\in\mathrm{D}(\mathcal{A}^{\frac{3}{2}})$ *and* $\boldsymbol{f}\in\mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))$. *Then, there exists a unique solution* $\boldsymbol{v}$ *to system*

(3) *in the sense of Definition 3.1. In addition,* $\boldsymbol{v}\in \mathrm{L}^p(\Omega;\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}})))$, $p\geq 2$, *and for any given* $\hat{c}>0$

$$\mathbb{E}\bigg[\exp\bigg\{\hat{c}\int_0^t\|\boldsymbol{v}(s)\|_{\mathbb{H}^3}^2\mathrm{d}s\bigg\}\bigg]<+\infty, \tag{33}$$

for all $t\in[0,T]$.

Remark 3.3. Note that we are considering deterministic initial data, that is, $\boldsymbol{v}_0\in \mathrm{D}(\mathcal{A}^{\frac{3}{2}})$, but one can also take $\boldsymbol{v}_0\in \mathrm{L}^2(\Omega;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))$ under an appropriate condition, which could help in obtaining (33).

3.1. **Proof of Theorem 3.2.** As discussed in introduction, the existence and uniqueness of a strong solution $\boldsymbol{v}$ to system (31) are established through the decomposition $\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{z}$, where $\boldsymbol{u}$ and $\boldsymbol{z}$ denote the unique solutions of systems (17) and (16), respectively.

We already know the existence of a unique regular solution to system (16), see Proposition 2.2. For $\boldsymbol{f}\in \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))$, $\mathrm{Tr}\left((-\Delta)^{3-2\gamma}GG^*\right)<\infty$, $\gamma\in(0,\frac12)$ and $\boldsymbol{v}_0\in\mathrm{D}(\mathcal{A}^{\frac{3}{2}})$ (see Proposition 2.2), we have for $p\geq 2$

$$\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{v}_0\|_2^2+\int_0^T\|\boldsymbol{f}(t)\|_{\dot{\mathbb{H}}^1}^2\mathrm{d}t+\sup_{t\in[0,T]}\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^p<+\infty,\ \ \mathbb{P}\text{-a.s.}, \tag{34}$$

that is, there exists a set $\widetilde{\Omega}\subset\Omega$ with $\mathbb{P}(\widetilde{\Omega})=1$ such that (34) holds for each $\omega\in\widetilde{\Omega}$.

Fix $T>0$ and $\omega\in\widetilde{\Omega}$. For each ω, we treat (17) as a deterministic problem with data $\boldsymbol{v}_0$, $\boldsymbol{f}$, and $\boldsymbol{z}$ satisfying (34), and establish the existence and uniqueness of its solution.

Taking projection $\mathcal{P}$ to system (17), we obtain

$$\begin{cases}\partial_t\Upsilon(\boldsymbol{u})+\nu\mathcal{A}\boldsymbol{u}+\mathcal{P}[((\boldsymbol{u}+\boldsymbol{z})\cdot\nabla)\Upsilon(\boldsymbol{u}+\boldsymbol{z})]+\mathcal{P}\bigg[\displaystyle\sum_{j=1}^2(\Upsilon(\boldsymbol{u}+\boldsymbol{z}))^j\nabla(\boldsymbol{u}+\boldsymbol{z})^j\bigg]\\ -(\alpha_1+\alpha_2)\mathcal{J}(\boldsymbol{u}+\boldsymbol{z})-\beta\mathcal{K}(\boldsymbol{u}+\boldsymbol{z})\\ =\mathcal{P}\boldsymbol{f}+\boldsymbol{z}+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}+\theta\alpha_1\mathcal{A}^2\boldsymbol{z},\quad \text{for a.e. } t\in(0,T),\\ \boldsymbol{u}(0)=\boldsymbol{v}_0,\end{cases} \tag{35}$$

in $\mathbb{H}$. We will show the existence of unique solution to system (35) in the following sense:

Definition 3.4. A function $\boldsymbol{u}\in\mathrm{C}([0,T];\mathrm{D}(\mathcal{A}))\cap\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac32}))$ with $\partial_t\boldsymbol{u}\in\mathrm{L}^2(0,T;\mathrm{D}(\mathcal{A}))$, is called a *strong solution* to system (35), if for $\boldsymbol{v}_0$, $\boldsymbol{f}$ and $\boldsymbol{z}$ satisfying (34), it satisfies

(i) for any $\psi\in\mathbb{H}$,

$$\begin{aligned}&(\Upsilon(\boldsymbol{u}(t)),\psi)\\ &=-\int_0^t\langle\nu\mathcal{A}\boldsymbol{u}(s)-(\alpha_1+\alpha_2)\mathcal{J}(\boldsymbol{u}(s)+\boldsymbol{z}(s))-\beta\mathcal{K}(\boldsymbol{u}(s)+\boldsymbol{z}(s)),\psi\rangle\mathrm{d}s\\ &\quad+\int_0^t b(\boldsymbol{u}(s)+\boldsymbol{z}(s),\psi(s),\Upsilon(\boldsymbol{u}(s)+\boldsymbol{z}(s)))\mathrm{d}s\\ &\quad-\int_0^t b(\psi(s),\boldsymbol{u}(s)+\boldsymbol{z}(s),\Upsilon(\boldsymbol{u}(s)+\boldsymbol{z}(s)))\mathrm{d}s\end{aligned}$$

$$+\int_0^t (\boldsymbol{f}(s)+\boldsymbol{z}(s)+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}(s)+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}(s),\psi)\mathrm{d}s, \tag{36}$$

for all $t\in[0,T]$;

(ii) the initial data:

$$\boldsymbol{u}(0)=\boldsymbol{v}_0 \ \text{ in } \ \mathrm{D}(\mathcal{A}^{\frac{3}{2}}).$$

Theorem 3.5. *There exists a unique solution $\boldsymbol{u}$ to system (35) in the sense of Definition 3.4.*

The proof of Theorem 3.5 is divided into the several steps which are given as separate subsubsections in the sequel.

3.1.1. *Finite-dimensional approximation.* Here, we consider the basis $\{\boldsymbol{w}_i\}_{i\in\mathbb{N}}$ of eigenfunctions of the Stokes operator $\mathcal{A}$, which is orthonormal in $\mathbb{H}$ and orthogonal in $\mathbb{V}$. We define $\mathbb{H}_n:=\operatorname{span}\{\boldsymbol{w}_j : j=1,\dots,n\}$ for $n\in\mathbb{N}$. Let P_n denote the orthogonal projection from $\mathbb{V}'$ onto $\mathbb{H}_n$. For every $\boldsymbol{x}\in\mathbb{V}'$, the projection is given by $\mathrm{P}_n\boldsymbol{x}=\sum\limits_{j=1}^{n}\langle\boldsymbol{x},\boldsymbol{w}_j\rangle\boldsymbol{w}_j$. Since each element $\boldsymbol{x}\in\mathbb{H}$ induces a functional $\boldsymbol{x}^*\in\mathbb{V}'$ defined by $\langle\boldsymbol{x}^*,\boldsymbol{y}\rangle=(\boldsymbol{x},\boldsymbol{y})$, $\boldsymbol{y}\in\mathbb{V}$, the restriction $\mathrm{P}_n|_{\mathbb{H}}$ corresponds to the orthogonal projection from $\mathbb{H}$ onto $\mathbb{H}_n$. Therefore, for all $\boldsymbol{x}\in\mathbb{H}$, $\mathrm{P}_n\boldsymbol{x}=\sum\limits_{i=1}^{n}(\boldsymbol{x},\boldsymbol{w}_i)\,\boldsymbol{w}_i$ and $\|\mathrm{P}_n\boldsymbol{x}-\boldsymbol{x}\|_2\to 0$ as $n\to\infty$. By the relation (9), for all $\boldsymbol{x}\in\mathbb{V}$, $\mathrm{P}_n\boldsymbol{x}$ also represents the orthogonal projection with respect to the inner product $(\cdot,\cdot)_{\mathbb{V}}$ in $\mathbb{V}$, implying that $\|\mathrm{P}_n\boldsymbol{x}-\boldsymbol{x}\|_{\mathbb{V}}\to 0$ as $n\to\infty$. For each $n\in\mathbb{N}$, we search for a approximate solution of the form

$$\boldsymbol{u}_n(x,t):=\sum_{k=1}^{n} g_k^n(t)\boldsymbol{w}_k(x),$$

where $g_1^n(t),\dots,g_k^n(t)$ are unknown scalar functions of t such that it satisfies the following finite-dimensional system of ordinary differential equations in $\mathbb{H}_n$:

$$\begin{cases}\partial_t\Upsilon(\boldsymbol{u}_n)+\nu\mathcal{A}\boldsymbol{u}_n+\mathrm{P}_n\mathcal{P}[((\boldsymbol{u}_n+\boldsymbol{z})\cdot\nabla)\Upsilon(\boldsymbol{u}_n+\boldsymbol{z})]\\[2pt] +\mathrm{P}_n\mathcal{P}\left[\displaystyle\sum_{j=1}^{2}(\Upsilon(\boldsymbol{u}_n+\boldsymbol{z}))^j\nabla(\boldsymbol{u}_n+\boldsymbol{z})^j\right]\\[2pt] -(\alpha_1+\alpha_2)\mathrm{P}_n[\mathcal{J}(\boldsymbol{u}_n+\boldsymbol{z})]-\beta\mathrm{P}_n[\mathcal{K}(\boldsymbol{u}_n+\boldsymbol{z})]\\[2pt] =\boldsymbol{f}_n+\boldsymbol{z}_n+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}_n+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}_n, \text{ for a.e. } t\in(0,T),\\ \boldsymbol{u}_n(0)=\mathrm{P}_n[\boldsymbol{v}_0]=:\boldsymbol{u}_{0,n},\end{cases} \tag{37}$$

where $\boldsymbol{f}_n:=\mathrm{P}_n[\mathcal{P}\boldsymbol{f}]$ and $\boldsymbol{z}_n:=\mathrm{P}_n[\boldsymbol{z}]$. So we have to solve a system of n ordinary differential equations (ODE). Classical results on ODE ensures the existence of g_k^n in an interval $[0,T_n]$, for $0<T_n\le T$. We obtain the time $T_n=T$ by establishing the uniform energy estimates of the solutions satisfied by the system (37).

3.1.2. *A priori estimates.* **Step I.** Taking the inner product in $(37)_1$ with $\boldsymbol{u}_n$, we get

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\boldsymbol{u}_n(t)\|_2^2+\alpha_1\|\nabla\boldsymbol{u}_n(t)\|_2^2\right]+\nu\|\nabla\boldsymbol{u}_n(t)\|_2^2+\frac{\beta}{2}\|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_4^4\\ &=\underbrace{\alpha_1 b(\boldsymbol{z}(t),\boldsymbol{u}_n(t),\mathcal{A}\boldsymbol{u}_n(t))}_{I_1}+\underbrace{b(\boldsymbol{z}(t),\boldsymbol{u}_n(t),\Upsilon(\boldsymbol{z}(t)))}_{I_2}-\underbrace{b(\boldsymbol{u}_n(t),\boldsymbol{z}(t),\boldsymbol{u}_n(t))}_{I_3}\end{aligned}$$

$$
\begin{aligned}
&- \underbrace{\alpha_1 b(\boldsymbol{u}_n(t), \boldsymbol{z}(t), \mathcal{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)))}_{I_4} \\
&- \underbrace{\frac{(\alpha_1+\alpha_2)}{2}\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))]^2 : \mathrm{A}(\boldsymbol{u}_n(x,t))\mathrm{d}x}_{I_5} \\
&+ \underbrace{\frac{\beta}{2}\int_{\mathbb{T}^2}|\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t)) : \mathrm{A}(\boldsymbol{z}(x,t))]\mathrm{d}x}_{I_6} \\
&+ \underbrace{[(\boldsymbol{f}(t), \boldsymbol{u}_n(t)) + (\boldsymbol{z}(t)+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}(t)+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}(t), \boldsymbol{u}_n(t))]}_{I_7}, \qquad (38)
\end{aligned}
$$

for a.e. $t \in [0,T]$. Let us estimate each term on the right hand side of (38) using integration by parts, and Hölder's, Sobolev and Young's inequalities as follows:

$$
\begin{aligned}
|I_1| &= |\alpha_1 b(\boldsymbol{z}, \boldsymbol{u}_n, \mathcal{A}\boldsymbol{u}_n)| \\
&= \left|\alpha_1 \int_{\mathbb{T}^2}\sum_{i,j,k}^{2}\partial_k \boldsymbol{z}^i \partial_i \boldsymbol{u}_n^j \partial_k \boldsymbol{u}_n^j \mathrm{d}x\right| \\
&\leq C\|\boldsymbol{z}\|_{\mathbb{W}^{1,\infty}}\|\nabla \boldsymbol{u}_n\|_2^2 \\
&\leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\|\nabla \boldsymbol{u}_n\|_2^2, \qquad (39)
\end{aligned}
$$

$$
|I_2| \leq \|\boldsymbol{z}\|_\infty \|\nabla \boldsymbol{u}_n\|_2 \|\boldsymbol{z}+\alpha_1\mathcal{A}\boldsymbol{z}\|_2 \leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^2\|\nabla \boldsymbol{u}_n\|_2 \leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4 + C\|\nabla \boldsymbol{u}_n\|_2^2, \quad (40)
$$

$$
|I_3| \leq \|\nabla \boldsymbol{z}\|_2 \|\boldsymbol{u}_n\|_4^2 \leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\|\nabla \boldsymbol{u}_n\|_2^2, \qquad (41)
$$

$$
\begin{aligned}
|I_4| &= \alpha_1\left|\int_{\mathbb{T}^2}\sum_{i,j,k=1}^{2}\partial_k \boldsymbol{u}_n^i \partial_i \boldsymbol{z}^j \partial_k(\boldsymbol{u}_n^j+\boldsymbol{z}^j)\mathrm{d}x + \int_{\mathbb{T}^2}\sum_{i,j,k=1}^{2}\boldsymbol{u}_n^i\partial_i\partial_k \boldsymbol{z}^j \partial_k(\boldsymbol{u}_n^j+\boldsymbol{z}^j)\mathrm{d}x\right| \\
&\leq C\|\nabla \boldsymbol{u}_n\|_2\|\nabla \boldsymbol{z}\|_4\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4 + C\|\boldsymbol{u}_n\|_4\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4 \\
&\leq C\|\nabla \boldsymbol{u}_n\|_2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4 \\
&\leq \frac{\beta}{12}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^4 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4 + C\|\nabla \boldsymbol{u}_n\|_2^2, \qquad (42)
\end{aligned}
$$

$$
|I_5| \leq \frac{\beta}{12}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^4 + C\|\nabla \boldsymbol{u}_n\|_2^2, \qquad (43)
$$

$$
|I_6| \leq \frac{\beta}{12}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^4 + C\|\mathrm{A}(\boldsymbol{z})\|_4^4 \leq \frac{\beta}{12}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^4 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4, \qquad (44)
$$

$$
|I_7| \leq C[\|\boldsymbol{f}\|_2 + \|\boldsymbol{z}\|_{\dot{\mathbb{H}}^4}]\|\boldsymbol{u}_n\|_2 \leq C[\|\boldsymbol{f}\|_2^2 + \|\boldsymbol{z}\|_{\dot{\mathbb{H}}^4}^4 + 1] + C\|\boldsymbol{u}_n\|_2^2. \qquad (45)
$$

Combining (38)-(45), we obtain

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\nabla \boldsymbol{u}_n(t)\|_2^2\right] + 2\nu\|\nabla \boldsymbol{u}_n(t)\|_2^2 + \frac{\beta}{2}\|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_4^4 \\
&\leq C[\|\boldsymbol{f}(t)\|_2^2 + \|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^4}^4 + 1] + C[1+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^3}]\left[\|\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\nabla \boldsymbol{u}_n(t)\|_2^2\right], \quad (46)
\end{aligned}
$$

for a.e. $t \in [0,T]$, which, by an application of the Gronwall Lemma, gives

$$
\|\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\nabla \boldsymbol{u}_n(t)\|_2^2 + 2\nu\int_0^t \|\nabla \boldsymbol{u}_n(s)\|_2^2\mathrm{d}s + \frac{\beta}{2}\int_0^t \|\mathrm{A}(\boldsymbol{u}_n(s)+\boldsymbol{z}(s))\|_4^4\mathrm{d}s
$$

$$
\begin{aligned}
&\leq \left[\|\boldsymbol{u}_n(0)\|_2^2 + \alpha_1\|\nabla \boldsymbol{u}_n(0)\|_2^2 + C\int_0^t \{\|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^4 + 1\}\mathrm{d}s\right] \\
&\quad \times e^{C\int_0^t\{1+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}\}\mathrm{d}s} \\
&\leq C\left[\|\boldsymbol{v}_0\|_{\mathbb{V}}^2 + \int_0^t \{\|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^4 + 1\}\mathrm{d}s\right] e^{C\int_0^t\{1+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}\}\mathrm{d}s} \\
&:= K_1(t),
\end{aligned} \tag{47}
$$

for all $t \in [0,T]$. In view of (34), we have from (47) that

$$
\{\boldsymbol{u}^n\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^\infty(0,T;\mathbb{V}) \cap \mathrm{L}^4(0,T;\mathbb{W}^{1,4}(\mathbb{T}^2)). \tag{48}
$$

Step II. Since $\{\boldsymbol{w}_i\}_{i\in\mathbb{N}}$ is a sequence of eigenfunctions of the Stokes operator $\mathcal{A}$, it gives $\mathcal{A}\boldsymbol{u}_n \in \mathbb{H}_n$. Taking the inner product in $(37)_1$ with $\mathcal{A}\boldsymbol{u}_n$, we get

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\nabla \boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2\right] + \nu\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2 \\
&+ \underbrace{\frac{\beta}{2}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t)) : \mathrm{A}(\mathcal{A}\boldsymbol{u}_n(x,t)+\mathcal{A}\boldsymbol{z}(x,t))]\mathrm{d}x}_{\widetilde{I}_1} \\
&= -\underbrace{\alpha_1 b(\mathcal{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)), \boldsymbol{u}_n(t)+\boldsymbol{z}(t), \mathcal{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)))}_{\widetilde{I}_2} \\
&\quad - \underbrace{b(\boldsymbol{z}(t)-\alpha_1\mathcal{A}\boldsymbol{z}(t), \boldsymbol{u}_n(t), \mathcal{A}\boldsymbol{u}_n(t))}_{\widetilde{I}_3} + \underbrace{\alpha_1 b(\mathcal{A}\boldsymbol{z}(t), \boldsymbol{u}_n(t)+\boldsymbol{z}(t), \mathcal{A}\boldsymbol{z}(t))}_{\widetilde{I}_4} \\
&\quad + \underbrace{\alpha_1 b(\mathcal{A}\boldsymbol{z}(t), \boldsymbol{z}(t), \mathcal{A}\boldsymbol{u}_n(t))}_{\widetilde{I}_5} - \underbrace{b(\boldsymbol{u}_n(t)+\boldsymbol{z}(t), \Upsilon(\boldsymbol{z}(t)), \mathcal{A}\boldsymbol{u}_n(t))}_{\widetilde{I}_6} \\
&\quad - \underbrace{\frac{(\alpha_1+\alpha_2)}{2}\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))]^2 : \mathrm{A}(\mathcal{A}\boldsymbol{u}_n(x,t))\mathrm{d}x}_{\widetilde{I}_7} \\
&\quad + \underbrace{\frac{\beta}{2}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t)) : \mathrm{A}(\mathcal{A}\boldsymbol{z}(x,t))]\mathrm{d}x}_{\widetilde{I}_8} \\
&\quad + \underbrace{[(\boldsymbol{f}(t), \mathcal{A}\boldsymbol{u}_n(t)) + (\boldsymbol{z}(t)+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}(t)+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}(t), \mathcal{A}\boldsymbol{u}_n(t))]}_{\widetilde{I}_9},
\end{aligned} \tag{49}
$$

for a.e. $t \in [0,T]$. Let us estimate each term on the right hand side of (49) using integration by parts, and Hölder's, Sobolev and Young's inequalities as follows:

$$
\begin{aligned}
\widetilde{I}_1 &= \frac{\beta}{2}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2 |\nabla \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\mathrm{d}x \\
&\quad + \frac{\beta}{4}\int_{\mathbb{T}^2} \left|\nabla\left(|\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\right)\right|^2\mathrm{d}x,
\end{aligned} \tag{50}
$$

$$
\begin{aligned}
|\widetilde{I}_2| &= \left|\frac{\alpha_1}{2}\int_{\mathbb{T}^2}[\mathcal{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))\cdot \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))]\cdot\mathcal{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))\mathrm{d}x\right| \\
&= \left|\frac{\alpha_1}{2}\int_{\mathbb{T}^2}[\mathrm{div}\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))\cdot \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))]\cdot\mathcal{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))\mathrm{d}x\right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\beta}{8}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2|\nabla \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\mathrm{d}x + C\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2^2\\
&\leq \frac{\beta}{8}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2|\nabla \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\mathrm{d}x + C\|\mathcal{A}\boldsymbol{u}_n\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2\\
&\leq \frac{\beta}{8}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2|\nabla \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\mathrm{d}x + C\big[\|\mathcal{A}\boldsymbol{u}_n\|_2^2 + \|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4 + 1\big],
\end{aligned} \tag{51}
$$

$$
|\widetilde{I}_3| \leq \|\boldsymbol{z}-\alpha_1\mathcal{A}\boldsymbol{z}\|_4\|\nabla\boldsymbol{u}_n\|_4\|\mathcal{A}\boldsymbol{u}_n\|_2 \leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\boldsymbol{u}_n\|_2^2, \tag{52}
$$

$$
|\widetilde{I}_4| \leq C\|\nabla(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\mathcal{A}\boldsymbol{z}\|_4^2 \leq C\|\nabla\boldsymbol{u}_n\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4 + C, \tag{53}
$$

$$
|\widetilde{I}_5| \leq C\|\mathcal{A}\boldsymbol{z}\|_2\|\nabla\boldsymbol{z}\|_\infty\|\mathcal{A}\boldsymbol{u}_n\|_2 \leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2\|\mathcal{A}\boldsymbol{u}_n\|_2 \leq C\|\mathcal{A}\boldsymbol{u}_n\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4, \tag{54}
$$

$$
\begin{aligned}
|\widetilde{I}_6| &\leq C\|\boldsymbol{u}_n+\boldsymbol{z}\|_\infty\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\boldsymbol{u}_n\|_2\\
&\leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\boldsymbol{u}_n\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2\|\mathcal{A}\boldsymbol{u}_n\|_2\\
&\leq C[1+\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}]\|\mathcal{A}\boldsymbol{u}_n\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4,
\end{aligned} \tag{55}
$$

$$
\begin{aligned}
|\widetilde{I}_7| &\leq C\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))||\nabla \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))||\mathcal{A}\boldsymbol{u}_n(x)|\mathrm{d}x\\
&\leq \frac{\beta}{4}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2|\nabla \mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\mathrm{d}x + C\|\mathcal{A}\boldsymbol{u}_n\|_2^2,
\end{aligned} \tag{56}
$$

$$
\begin{aligned}
|\widetilde{I}_8| &\leq C\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_{12}\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_3\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq C\||\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})|^2\|_6^{\frac12}\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq C\||\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})|^2\|_{\dot{\mathbb{H}}^1}^{\frac12}\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq C\|\mathrm{A}(\boldsymbol{u}_n)\|_4\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\quad + C\left(\int_{\mathbb{T}^2} |\nabla\left(|\mathrm{A}(\boldsymbol{u}_n(x))|^2\right)|^2\mathrm{d}x\right)^{\frac14}\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq \frac{\beta}{8}\int_{\mathbb{T}^2} |\nabla\left(|\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\right)|^2\mathrm{d}x + C\|\mathrm{A}(\boldsymbol{u}_n)\|_4^4 + C\|\mathcal{A}\boldsymbol{u}_n\|_2^2\\
&\quad + C\|\mathcal{A}\boldsymbol{z}\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4\\
&\leq \frac{\beta}{8}\int_{\mathbb{T}^2} |\nabla\left(|\mathrm{A}(\boldsymbol{u}_n(x)+\boldsymbol{z}(x))|^2\right)|^2\mathrm{d}x + C\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^4\\
&\quad + C\|\mathcal{A}\boldsymbol{u}_n\|_2^2 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^4 + C,
\end{aligned} \tag{57}
$$

$$
|\widetilde{I}_9| \leq C[\|\boldsymbol{f}\|_2 + \|\boldsymbol{z}\|_{\dot{\mathbb{H}}^4}]\|\mathcal{A}\boldsymbol{u}_n\|_2 \leq C[\|\boldsymbol{f}\|_2^2 + \|\boldsymbol{z}\|_{\dot{\mathbb{H}}^4}^4 + 1] + C\|\mathcal{A}\boldsymbol{u}_n\|_2^2. \tag{58}
$$

Combining (49)-(58), we obtain

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\nabla\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2\right] + 2\nu\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2\\
&+ \frac{\beta}{2}\int_{\mathbb{T}^2} |\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2|\nabla \mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2\mathrm{d}x\\
&+ \frac{\beta}{4}\int_{\mathbb{T}^2} |\nabla\left(|\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2\right)|^2\mathrm{d}x\\
&\leq C[\|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_4^4 + \|\boldsymbol{f}(t)\|_2^2 + \|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^4}^4 + 1]\\
&\quad + C[1+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^3}]\big[\|\nabla\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2\big],
\end{aligned} \tag{59}
$$

for a.e. $t \in [0,T]$, which, by an application of the Gronwall Lemma, gives

$$
\begin{aligned}
&\|\nabla \boldsymbol{u}_n(t)\|_2^2 + \alpha_1 \|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2 + 2\nu \int_0^t \|\mathcal{A}\boldsymbol{u}_n(s)\|_2^2 \mathrm{d}s \\
&+ \frac{\beta}{4}\int_0^t \int_{\mathbb{T}^2} |\nabla \left(|\mathrm{A}(\boldsymbol{u}_n(x,s)+\boldsymbol{z}(x,s))|^2\right)|^2 \mathrm{d}x\mathrm{d}s \\
&\leq \Bigg[\|\nabla \boldsymbol{u}_n(0)\|_2^2 + \alpha_1 \|\mathcal{A}\boldsymbol{u}_n(0)\|_2^2 \\
&\quad + C\int_0^t \{\|\mathrm{A}(\boldsymbol{u}_n(s)+\boldsymbol{z}(s))\|_4^4 + \|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^4 + 1\}\mathrm{d}s\Bigg] e^{C\int_0^t \{1+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}\}\mathrm{d}s} \\
&\leq C\Bigg[\|\mathcal{A}\boldsymbol{v}_0\|_2^2 + \int_0^t \{\|\mathrm{A}(\boldsymbol{u}_n(s)+\boldsymbol{z}(s))\|_4^4 + \|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^4 + 1\}\mathrm{d}s\Bigg] \\
&\quad \times e^{C\int_0^t \{1+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}\}\mathrm{d}s} \\
&\leq C\Bigg[\|\mathcal{A}\boldsymbol{v}_0\|_2^2 + K_1(t) + \int_0^t \{\|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^4 + 1\}\mathrm{d}s\Bigg] e^{C\int_0^t \{1+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}\}\mathrm{d}s} \\
&:= K_2(t),
\end{aligned} \tag{60}
$$

for all $t \in [0,T]$. In view of (34) and (47), we have from (60) that

$$
\{\boldsymbol{u}^n\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A})). \tag{61}
$$

Step III. Taking the inner product with $\partial_t \boldsymbol{u}_n$ in $(37)_1$, we obtain

$$
\begin{aligned}
&\|\partial_t \boldsymbol{u}_n(t)\|_2^2 + \alpha_1 \|\partial_t \nabla \boldsymbol{u}_n(t)\|_2^2 + \frac{\nu}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla \boldsymbol{u}_n(t)\|_2^2 \\
&= \underbrace{b(\boldsymbol{u}_n(t)+\boldsymbol{z}(t), \partial_t \boldsymbol{u}_n(t), \Upsilon(\boldsymbol{u}_n(t)+\boldsymbol{z}_n(t)))}_{\widetilde{I}_1} \\
&\quad - \underbrace{b(\partial_t \boldsymbol{u}_n(t), \boldsymbol{u}_n(t)+\boldsymbol{z}(t), \Upsilon(\boldsymbol{u}_n(t)+\boldsymbol{z}_n(t)))}_{\widetilde{I}_2} \\
&\quad - \underbrace{(\alpha_1+\alpha_2)(\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))^2, \partial_t \nabla \boldsymbol{u}_n(t))}_{\widetilde{I}_3} \\
&\quad - \underbrace{\beta(|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))|^2 \mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)), \partial_t \nabla \boldsymbol{u}_n(t))}_{\widetilde{I}_4} \\
&\quad + \underbrace{(\boldsymbol{f}(t)+\boldsymbol{z}(t)+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}(t)+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}(t), \partial_t \boldsymbol{u}_n(t))}_{\widetilde{I}_5},
\end{aligned} \tag{62}
$$

for a.e. $t \in [0,T]$. Let us estimate each term on the right hand side of (62) using integration by parts, and Hölder's, Sobolev and Young's inequalities as follows:

$$
\begin{aligned}
|\widetilde{I}_1| &\leq \|\boldsymbol{u}_n+\boldsymbol{z}\|_\infty \|\partial_t \nabla \boldsymbol{u}_n\|_2 \|\boldsymbol{u}_n + \alpha_1 \mathcal{A}\boldsymbol{u}_n + \boldsymbol{z} + \alpha_1 \mathcal{A}\boldsymbol{z}\|_2 \\
&\leq C\|\partial_t \nabla \boldsymbol{u}_n\|_2 \|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2^2 \\
&\leq \frac{\alpha_1}{8}\|\partial_t \nabla \boldsymbol{u}_n\|_2^2 + C\|\mathcal{A}\boldsymbol{u}_n\|_2^6 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^6 + C,
\end{aligned} \tag{63}
$$

$$
\begin{aligned}
|\widetilde{I}_2| &\leq \|\partial_t \boldsymbol{u}_n\|_4 \|\nabla(\boldsymbol{u}_n+\boldsymbol{z})\|_4 \|\boldsymbol{u}_n + \alpha_1 \mathcal{A}\boldsymbol{u}_n + \boldsymbol{z} + \alpha_1 \mathcal{A}\boldsymbol{z}\|_2 \\
&\leq C\|\partial_t \nabla \boldsymbol{u}_n\|_2 \|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2^2
\end{aligned}
$$

$$\leq \frac{\alpha_1}{8}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathcal{A}\boldsymbol{u}_n\|_2^6 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^6 + C, \tag{64}$$

$$\begin{aligned}
|\widetilde{I}_3| &\leq \|\partial_t\nabla\boldsymbol{u}_n\|_2\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^2 \\
&\leq \frac{1}{2}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_4^4 \\
&\leq \frac{\alpha_1}{8}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2^4 \\
&\leq \frac{1}{2}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathcal{A}\boldsymbol{u}_n\|_2^6 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^6 + C,
\end{aligned} \tag{65}$$

$$\begin{aligned}
|\widetilde{I}_4| &\leq \|\partial_t\nabla\boldsymbol{u}_n\|_2\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_6^3 \\
&\leq \frac{\alpha_1}{8}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_6^6 \\
&\leq \frac{1}{2}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathcal{A}(\boldsymbol{u}_n+\boldsymbol{z})\|_2^6 \\
&\leq \frac{1}{2}\|\partial_t\nabla\boldsymbol{u}_n\|_2^2 + C\|\mathcal{A}\boldsymbol{u}_n\|_2^6 + C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^6 + C,
\end{aligned} \tag{66}$$

$$|\widetilde{I}_5| \leq C[\|\boldsymbol{f}\|_2^2 + \|\boldsymbol{z}\|_{\dot{\mathbb{H}}^4}^6 + 1] + \frac{1}{2}\|\partial_t\boldsymbol{u}_n\|_2^2. \tag{67}$$

Combining (62)-(67), we obtain

$$\begin{aligned}
&\|\partial_t\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\partial_t\nabla\boldsymbol{u}_n(t)\|_2^2 + \nu\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla\boldsymbol{u}_n(t)\|_2^2 \\
&\leq C[\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^6 + \|\boldsymbol{f}(t)\|_2^2 + \|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^4}^6 + 1],
\end{aligned} \tag{68}$$

for a.e. $t \in [0,T]$, which, using (60), gives

$$\begin{aligned}
&\nu\|\nabla\boldsymbol{u}_n(t)\|_2^2 + \int_0^t \left[\|\partial_t\boldsymbol{u}_n(s)\|_2^2 + \alpha_1\|\partial_t\nabla\boldsymbol{u}_n(s)\|_2^2\right]\mathrm{d}s \\
&\leq \nu\|\nabla\boldsymbol{u}_n(0)\|_2^2 + C\int_0^t \{\|\mathcal{A}\boldsymbol{u}_n(s)\|_2^6 + \|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^6 + 1\}\mathrm{d}s \\
&\leq C\left[\|\boldsymbol{u}_0\|_{\mathbb{V}}^2 + t[K_2(t)]^3 + \int_0^t \{\|\boldsymbol{f}(s)\|_2^2 + \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^4}^6 + 1\}\mathrm{d}s\right] \\
&:= K_3(t),
\end{aligned} \tag{69}$$

for all $t \in [0,T]$. In view of (34) and (60), we have from (69) that

$$\{\partial_t\boldsymbol{u}^n\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^2(0,T;\mathbb{V}). \tag{70}$$

Let us define

$$\begin{aligned}
\mathcal{G}(\boldsymbol{u},\boldsymbol{z}) :=\ &\nu\mathcal{A}\boldsymbol{u} + \mathcal{P}[((\boldsymbol{u}+\boldsymbol{z})\cdot\nabla)\Upsilon(\boldsymbol{u}+\boldsymbol{z})] + \mathcal{P}\left[\sum_{j=1}^{2}[\Upsilon(\boldsymbol{u}+\boldsymbol{z})]^j\nabla(\boldsymbol{u}^j+\boldsymbol{z}^j)\right] \\
&- (\alpha_1+\alpha_2)\mathcal{P}\big[\mathrm{div}[\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})^2]\big] - \beta\mathcal{P}\big[\mathrm{div}[|\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})|^2\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})]\big].
\end{aligned} \tag{71}$$

Note that, from (37), we have

$$\begin{aligned}
&\int_0^T \|\mathrm{P}_n\mathcal{G}(\boldsymbol{u}_n(s),\boldsymbol{z}(s))\|_{\mathbb{V}'}^2\mathrm{d}s \\
&\leq \int_0^T \|\partial_t\boldsymbol{u}_n(s) + \alpha_1\partial_t\mathcal{A}\boldsymbol{u}_n(s) - \mathrm{P}_n\mathcal{P}\boldsymbol{f}(s) - \mathrm{P}_n\boldsymbol{z}(s) \\
&\quad - (\theta+\alpha_1-\nu)\mathcal{A}\mathrm{P}_n\boldsymbol{z}(s) - \theta\alpha_1\mathrm{P}_n\mathcal{A}^2\boldsymbol{z}(s)\|_{\mathbb{V}'}^2\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&\leq C\int_0^T \|\partial_t\nabla \boldsymbol{u}_n(s)\|_2^2\mathrm{d}s + C\int_0^T \|\boldsymbol{f}(s)\|_2^2\mathrm{d}s + C\int_0^T \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\\
&\leq CK_3(T) + C\int_0^T \|\boldsymbol{f}(s)\|_2^2\mathrm{d}s + C\int_0^T \|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s, \qquad (72)
\end{aligned}
$$

and (69) provides

$$
\begin{aligned}
\int_0^T \|\partial_t\Upsilon(\boldsymbol{u}_n(s))\|_{\mathbb{V}'}^2\mathrm{d}s &\leq 2\int_0^T \|\partial_t\boldsymbol{u}_n(s)\|_{\mathbb{V}'}^2\mathrm{d}s + 2\alpha_1^2\int_0^T \|\partial_t\mathcal{A}\boldsymbol{u}_n(s)\|_{\mathbb{V}'}^2\mathrm{d}s\\
&\leq C\int_0^T \|\partial_t\nabla\boldsymbol{u}_n(s)\|_2^2\mathrm{d}s\\
&\leq CK_3(T), \qquad (73)
\end{aligned}
$$

which implies

$$
\{\mathrm{P}_n\mathcal{G}(\boldsymbol{u}_n,\boldsymbol{z})\}_{n\in\mathbb{N}} \text{ and } \{\partial_t\Upsilon(\boldsymbol{u}_n)\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^2(0,T;\mathbb{V}'),\ \mathbb{P}\text{-a.s.} \qquad (74)
$$

3.1.3. *Weak*, weak and strong limits.* Using (48), (61), (70), (74) and the *Banach-Alaoglu theorem*, we infer the existence of an element $\boldsymbol{u}\in \mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$ with $\partial_t(\mathrm{I}+\alpha_1\mathcal{A})\boldsymbol{u}\in\mathrm{L}^2(0,T;\mathbb{V}')$ and $\mathcal{G}_0\in\mathrm{L}^2(0,T;\mathbb{V}')$ such that

$$
\left.
\begin{aligned}
\boldsymbol{u}_n &\xrightharpoonup{w^*} \boldsymbol{u} &&\text{in} && \mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A})),\\
\boldsymbol{u}_n &\xrightharpoonup{w} \boldsymbol{u} &&\text{in} && \mathrm{L}^2(0,T;\mathrm{D}(\mathcal{A}))\\
\partial_t\Upsilon(\boldsymbol{u}_n) &\xrightharpoonup{w} \partial_t\Upsilon(\boldsymbol{u}) &&\text{in} && \mathrm{L}^2(0,T;\mathbb{V}'),\\
\mathrm{P}_n\mathcal{G}(\boldsymbol{u}_n,\boldsymbol{z}) &\xrightharpoonup{w} \mathcal{G}_0 &&\text{in} && \mathrm{L}^2(0,T;\mathbb{V}'),
\end{aligned}
\right\} \qquad (75)
$$

along a subsequence (still denoted by the same symbol). In addition, since $\boldsymbol{u}_n\in \mathrm{L}^2(0,T;\mathrm{D}(\mathcal{A}))$, $\partial_t\boldsymbol{u}_n\in\mathrm{L}^2(0,T;\mathbb{H})$, in view of continuous embeddings $\mathrm{D}(\mathcal{A})\subset\mathbb{V}\subset\mathbb{H}$ with compact embedding $\mathrm{D}(\mathcal{A})\subset\mathbb{V}$ and Aubin-Lions compactness lemma, we also have (along a subsequence, still denoted by the same symbol)

$$
\boldsymbol{u}_n \to \boldsymbol{u} \quad \text{in} \quad \mathrm{L}^2(0,T;\mathbb{V}). \qquad (76)
$$

3.1.4. *Passing $n\to\infty$ in* (37). Let us start by showing that $\mathcal{G}(\boldsymbol{u},\boldsymbol{z})$ is an element in $\mathrm{L}^2(0,T;\mathbb{V}')$ for $\boldsymbol{u}\in\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$, $\boldsymbol{f}\in\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))$ and $\boldsymbol{z}\in\mathrm{L}^\infty(0,T;\dot{\mathbb{H}}^3(\mathbb{T}^2))$. Since, for $\boldsymbol{\phi}\in\mathrm{L}^2(0,T;\mathbb{V})$, we have

$$
\begin{aligned}
&\int_0^T \langle\mathcal{G}(\boldsymbol{u}(t),\boldsymbol{z}(t)),\boldsymbol{\phi}(t)\rangle\,\mathrm{d}t\\
&=\int_0^T\int_{\mathbb{T}^2}\Big\{\nu\nabla\boldsymbol{u}(t) + ((\boldsymbol{u}(t)+\boldsymbol{z}(t))\otimes(\Upsilon(\boldsymbol{u}(t)+\boldsymbol{z}(t))))\\
&\quad-(\alpha_1+\alpha_2)\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))^2 - \beta|\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))|^2\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\Big\} : \nabla\boldsymbol{\phi}(t)\mathrm{d}x\mathrm{d}t\\
&\quad+\int_0^T b(\boldsymbol{\phi}(t),\boldsymbol{u}(t)+\boldsymbol{z}(t),\Upsilon(\boldsymbol{u}(t)+\boldsymbol{z}(t)))\mathrm{d}t\\
&\leq C\int_0^T\Big[\|\boldsymbol{u}(t)\|_{\mathbb{V}} + \|\boldsymbol{u}(t)+\boldsymbol{z}(t)\|_\infty\|\Upsilon(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_2 + \|\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_4^2\\
&\quad+\|\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_6^3\Big]\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^T \|\boldsymbol{\phi}(t)\|_4\|\nabla(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_4\|\Upsilon(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_2\mathrm{d}t\\
&\leq C\int_0^T\left[\|\boldsymbol{u}(t)\|_{\mathbb{V}}+\|\mathcal{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_2^2+\|\mathcal{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_2^3\right]\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}\mathrm{d}t\\
&\leq C\left(\int_0^T\left[\|\boldsymbol{u}(t)\|_{\mathbb{V}}^2+\|\mathcal{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_2^4+\|\mathcal{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_2^6\right]\mathrm{d}t\right)^{\frac{1}{2}}\|\boldsymbol{\phi}\|_{\mathrm{L}^2(0,T;\mathbb{V})},
\end{aligned}
\tag{77}
$$

it implies that $\mathcal{G}(\boldsymbol{u},\boldsymbol{z})\in\mathrm{L}^2(0,T;\mathbb{V}')$.

Let us now show that $\mathcal{G}_0=\mathcal{G}(\boldsymbol{u},\boldsymbol{z})$. For $\boldsymbol{\phi}\in\mathrm{L}^2(0,T;\mathbb{V})$, we consider

$$
\begin{aligned}
&\int_0^T\langle\mathcal{G}_n(\boldsymbol{u}_n(t),\boldsymbol{z}(t))-\mathcal{G}(\boldsymbol{u}(t),\boldsymbol{z}(t)),\boldsymbol{\phi}(t)\rangle\,\mathrm{d}t\\
&=\underbrace{\int_0^T\langle\mathcal{G}(\boldsymbol{u}_n(t),\boldsymbol{z}(t)),\mathrm{P}_n\boldsymbol{\phi}(t)-\boldsymbol{\phi}(t)\rangle\,\mathrm{d}t}_{:=G_1(n)}\\
&\quad+\underbrace{\int_0^T\langle\mathcal{G}(\boldsymbol{u}_n(t),\boldsymbol{z}(t))-\mathcal{G}(\boldsymbol{u}(t),\boldsymbol{z}(t)),\boldsymbol{\phi}(t)\rangle\,\mathrm{d}t}_{:=G_2(n)}.
\end{aligned}
\tag{78}
$$

Since $\{\boldsymbol{u}_n\}_{n\in\mathbb{N}}$ is a bounded sequence in $\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$ (see (61)), a calculation similar to (77) gives that the sequence

$$\left\{\int_0^T\|\mathcal{G}(\boldsymbol{u}_n(t),\boldsymbol{z}(t))\|_{\mathbb{V}'}^2\mathrm{d}t\right\}_{n\in\mathbb{N}}\text{ is a bounded sequence.}$$

Also, we have $\mathrm{P}_n\boldsymbol{\phi}\to\boldsymbol{\phi}$ in $\mathrm{L}^2(0,T;\mathbb{V})$ therefore $\lim\limits_{n\to\infty}G_1(n)=0$.

Let us now show that $\lim\limits_{n\to\infty}G_2(n)=0$. Define $\boldsymbol{v}^n:=\boldsymbol{u}_n-\boldsymbol{u}$ and consider for $\boldsymbol{\phi}\in\mathrm{L}^2(0,T;\mathbb{V})$

$$
\begin{aligned}
&\langle\mathcal{G}(\boldsymbol{u}_n,\boldsymbol{z})-\mathcal{G}(\boldsymbol{u},\boldsymbol{z}),\boldsymbol{\phi}\rangle\\
&=\nu\underbrace{(\nabla\boldsymbol{v}^n,\nabla\boldsymbol{\phi})}_{G_{21}}-\underbrace{b(\boldsymbol{v}^n,\boldsymbol{\phi},\Upsilon(\boldsymbol{u}_n+\boldsymbol{z}))}_{G_{22}}-\underbrace{b(\boldsymbol{u}+\boldsymbol{z},\boldsymbol{\phi},\Upsilon(\boldsymbol{v}^n))}_{G_{23}}+\underbrace{b(\boldsymbol{\phi},\boldsymbol{u}+\boldsymbol{z},\Upsilon(\boldsymbol{v}^n))}_{G_{24}}+\underbrace{b(\boldsymbol{\phi},\boldsymbol{v}^n,\Upsilon(\boldsymbol{u}_n+\boldsymbol{z}))}_{G_{25}}\\
&\quad+(\alpha_1+\alpha_2)\underbrace{\left[(\mathrm{A}(\boldsymbol{v}^n)\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z}),\nabla\boldsymbol{\phi})+(\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})\mathrm{A}(\boldsymbol{v}^n),\nabla\boldsymbol{\phi})\right]}_{G_{26}}\\
&\quad+\beta\underbrace{\left[([\mathrm{A}(\boldsymbol{v}^n):\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})]\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})+[\mathrm{A}(\boldsymbol{u}+\boldsymbol{z}):\mathrm{A}(\boldsymbol{v}^n)]\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})+|\mathrm{A}(\boldsymbol{u}+\boldsymbol{z})|^2\mathrm{A}(\boldsymbol{v}^n),\nabla\boldsymbol{\phi})\right]}_{G_{27}}.
\end{aligned}
\tag{79}
$$

It is immediate from the weak convergence $(75)_2$ that

$$\int_0^T[\nu G_{21}(t)-G_{23}(t)+G_{24}(t)]\mathrm{d}t\to0\ \text{ as }n\to\infty.\tag{80}$$

Next using Hölder's, Agmon's and Sobolev's inequalities, and the following Gagliardo-Nirenberg inequality for $1\leq s<\infty$

$$\|\nabla\boldsymbol{y}(t)\|_s\leq C\|\mathcal{A}(\boldsymbol{y}(t))\|_2^{1-\frac{1}{s}}\|\boldsymbol{y}(t)\|_2^{\frac{1}{s}},\ \text{for all }\boldsymbol{y}\in\mathrm{D}(\mathcal{A}),\tag{81}$$

we have

$$\begin{aligned}
&\left|\int_0^T [-G_{22}(t)+G_{25}(t)]\mathrm{d}t\right| \\
&\leq \int_0^T [\|\boldsymbol{v}^n(t)\|_\infty\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}+\|\boldsymbol{\phi}(t)\|_4\|\nabla\boldsymbol{v}^n(t)\|_4]\|\Upsilon(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_2\mathrm{d}t \\
&\leq C\int_0^T \left[\|\mathcal{A}\boldsymbol{v}^n(t)\|_2^{\frac{1}{2}}\|\boldsymbol{v}^n(t)\|_2^{\frac{1}{2}}+\|\mathcal{A}\boldsymbol{v}^n(t)\|_2^{\frac{3}{4}}\|\boldsymbol{v}^n(t)\|_2^{\frac{1}{4}}\right] \\
&\quad\times\left[\|\mathcal{A}\boldsymbol{u}_n(t)\|_2+\|\mathcal{A}\boldsymbol{z}(t)\|_2\right]\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}\mathrm{d}t \\
&\leq C\int_0^T \left[\|\mathcal{A}\boldsymbol{u}_n(t)\|_2+\|\mathcal{A}\boldsymbol{z}(t)\|_2\right]\|\mathcal{A}\boldsymbol{v}^n(t)\|_2^{\frac{3}{4}}\|\boldsymbol{v}^n(t)\|_2^{\frac{1}{4}}\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}\mathrm{d}t \\
&\leq C\sup_{t\in[0,T]}\left[\|\mathcal{A}\boldsymbol{u}_n(t)\|_2+\|\mathcal{A}\boldsymbol{z}(t)\|_2\right]\|\mathcal{A}\boldsymbol{v}^n\|_{\mathrm{L}^2(0,T;\mathbb{H})}^{\frac{3}{4}}\|\boldsymbol{v}^n\|_{\mathrm{L}^2(0,T;\mathbb{H})}^{\frac{1}{4}}\|\boldsymbol{\phi}\|_{\mathrm{L}^2(0,T;\mathbb{V})} \\
&\leq C\sup_{t\in[0,T]}\left[[K_2(t)]^{\frac{1}{2}}+\|\mathcal{A}\boldsymbol{z}(t)\|_2\right][K_2(T)]^{\frac{3}{8}}\|\boldsymbol{v}^n\|_{\mathrm{L}^2(0,T;\mathbb{H})}^{\frac{1}{4}}\|\boldsymbol{\phi}\|_{\mathrm{L}^2(0,T;\mathbb{V})} \\
&\to 0 \text{ as } n\to\infty, \qquad (82)
\end{aligned}$$

where we have also used the strong convergence obtained in (76). Finally, using Hölder's inequality and (81), we obtain

$$\begin{aligned}
&\left|\int_0^T [(\alpha_1+\alpha_2)G_{26}(t)+\beta G_{27}(t)]\mathrm{d}t\right| \\
&\leq C\int_0^T \Big[\|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_3+\|\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_3+\|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_6^2 \\
&\quad+\|\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_6\|\mathrm{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t))\|_6+\|\mathrm{A}(\boldsymbol{u}(t)+\boldsymbol{z}(t))\|_6^2\Big] \\
&\quad\times\|\mathrm{A}(\boldsymbol{v}^n(t))\|_6\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}\mathrm{d}t \\
&\leq C\int_0^T \left[1+\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2+\|\mathcal{A}\boldsymbol{z}(t)\|_2^2\right]\|\mathcal{A}\boldsymbol{v}^n(t)\|_2^{\frac{5}{6}}\|\boldsymbol{v}^n(t)\|_2^{\frac{1}{6}}\|\boldsymbol{\phi}(t)\|_{\mathbb{V}}\mathrm{d}t \\
&\leq C\sup_{t\in[0,T]}\left[1+K_2(t)+\|\mathcal{A}\boldsymbol{z}(t)\|_2^2\right][K_2(T)]^{\frac{5}{12}}\|\boldsymbol{v}^n\|_{\mathrm{L}^2(0,T;\mathbb{H})}^{\frac{1}{6}}\|\boldsymbol{\phi}\|_{\mathrm{L}^2(0,T;\mathbb{V})} \\
&\to 0 \text{ as } n\to\infty, \qquad (83)
\end{aligned}$$

where we have also used the strong convergence obtained in (76). Combing (79)-(80) and (82)-(83), we infer that $\lim\limits_{n\to\infty} G_2(n)=0$. Therefore, from $(75)_4$, (78) and uniqueness of weak limit, we obtain $\mathcal{G}_0=\mathcal{G}(\boldsymbol{u},\boldsymbol{z})$. Also, we have $\boldsymbol{f}_n\to\mathcal{P}\boldsymbol{f}$ in $\mathrm{L}^2(0,T;\mathbb{H})$. Therefore, on passing to limit as $n\to\infty$ in (37), the limit $\boldsymbol{u}(\cdot)$ satisfies:

$$\partial_t(\mathrm{I}+\alpha_1\mathcal{A})\boldsymbol{u}=-\mathcal{G}(\boldsymbol{u},\boldsymbol{z})+\mathcal{P}\boldsymbol{f}+\boldsymbol{z}+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}, \quad \text{in } \mathrm{L}^2(0,T;\mathbb{V}'). \qquad (84)$$

Next, we notice that [31, Lemma 1.2, p.176], $(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}\in\mathrm{L}^2(0,T;\mathbb{V})$ and $\partial_t(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}\in\mathrm{L}^2(0,T;\mathbb{V}')$ imply $(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}\in\mathrm{C}([0,T];\mathbb{H})$, the real-valued function $t\mapsto\|(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}(t)\|_2^2$ is absolutely continuous and the following equality is satisfied:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}(t)\|_2^2=\left\langle\partial_t(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}(t),(\mathrm{I}+\alpha_1\mathcal{A})^{\frac{1}{2}}\boldsymbol{u}(t)\right\rangle$$

$$= \langle \partial_t(\mathrm{I} + \alpha_1\mathcal{A})\boldsymbol{u}(t), \boldsymbol{u}(t)\rangle , \tag{85}$$

for a.e. $t \in [0, T]$. Hence, the initial condition in the Definition 3.4 also makes sense.

3.1.5. *Uniqueness.* In order to obtain the uniqueness, we recall the following identity (form [6, Appendix]) for any vector $\boldsymbol{w}$ which will be used in the sequel:

$$\begin{aligned}
&\frac{1}{2}\nabla\left[\alpha_1|\nabla\boldsymbol{w}|^2 - |\boldsymbol{w}|^2\right] \\
&= (\boldsymbol{w}\cdot\nabla)\boldsymbol{w} + \mathrm{div}\big[-\alpha_1 \underbrace{\{\boldsymbol{w}\cdot\nabla\mathrm{A}(\boldsymbol{w}) + L(\boldsymbol{w})^T\mathrm{A}(\boldsymbol{w}) + \mathrm{A}(\boldsymbol{w})L(\boldsymbol{w})\}}_{=:\mathcal{N}(\boldsymbol{w})} - \alpha_2\mathrm{A}(\boldsymbol{w})^2\big] \\
&\quad - \underbrace{\left[(\boldsymbol{w}\cdot\nabla)(\boldsymbol{w} - \alpha_1\Delta\boldsymbol{w}) + \sum_{j=1}^{2}(\boldsymbol{w} - \alpha_1\Delta\boldsymbol{w})_j\nabla\boldsymbol{w}_j - (\alpha_1+\alpha_2)\mathrm{div}\mathrm{A}(\boldsymbol{w})^2\right]}_{=:\mathcal{M}(\boldsymbol{w})},
\end{aligned} \tag{86}$$

where $L(\boldsymbol{w})$ and $\mathrm{A}(\boldsymbol{w})$ are the matrix with entries $[L(\boldsymbol{w})]_{ij} = \partial_j\boldsymbol{w}_i$ and $[\mathrm{A}(\boldsymbol{w})]_{ij} = \partial_i\boldsymbol{w}_j + \partial_j\boldsymbol{w}_i$. We also recall Sobolev embedding (from [9, p. 723-724]) which says that there exists a constant $C_* > 0$ (independent of s) such that for any $\boldsymbol{w} \in \dot{\mathbb{H}}^1(\mathbb{T}^2)$

$$\|\boldsymbol{w}\|_s \leq C_* s\|\boldsymbol{w}\|_{\dot{\mathbb{H}}^1}. \tag{87}$$

Define $\mathfrak{U} = \boldsymbol{u}_1 - \boldsymbol{u}_2$, where $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are two solutions of system (17) in the sense of Definition 3.4. Then $\mathfrak{U} \in \mathrm{C}([0, T]; \mathbb{V}) \cap \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A}))$ and satisfies

$$\begin{cases}\partial_t\Upsilon(\mathfrak{U}(t)) = -\left[\mathcal{G}(\boldsymbol{u}_1(t), \boldsymbol{z}(t)) - \mathcal{G}(\boldsymbol{u}_2(t), \boldsymbol{z}(t))\right], \\ \mathfrak{U}(0) = \boldsymbol{0},\end{cases} \tag{88}$$

for a.e. $t \in [0, T]$, in $\mathbb{V}'$. Taking the inner product to the equation $(88)_1$ with $\mathfrak{U}$ and using integration by parts, [6, Equation (27)], (86), (11) and [6, Equation (30)], we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\mathfrak{U}\|_2^2 + \alpha_1\|\nabla\mathfrak{U}\|_2^2\right] + \nu\|\nabla\mathfrak{U}\|_2^2 \\
&+ \frac{\beta}{2}\int_{\mathbb{T}^2}|\mathrm{A}(\mathfrak{U}(x))|^2(|\mathrm{A}(\boldsymbol{u}_1(x) + \boldsymbol{z}(x))|^2 + |\mathrm{A}(\boldsymbol{u}_2(x) + \boldsymbol{z}(x))|^2)\mathrm{d}x \\
&+ \frac{\beta}{2}\int_{\mathbb{T}^2}(|\mathrm{A}(\boldsymbol{u}_1(x) + \boldsymbol{z}(x))|^2 - |\mathrm{A}(\boldsymbol{u}_2(x) + \boldsymbol{z}(x))|^2)^2\mathrm{d}x \\
&= -\langle\mathcal{M}(\boldsymbol{u}_1 + \boldsymbol{z}) - \mathcal{M}(\boldsymbol{u}_2 + \boldsymbol{z}), \mathfrak{U}\rangle \\
&= -\langle((\boldsymbol{u}_1 + \boldsymbol{z})\cdot\nabla)(\boldsymbol{u}_1 + \boldsymbol{z}) - ((\boldsymbol{u}_2 + \boldsymbol{z})\cdot\nabla)(\boldsymbol{u}_2 + \boldsymbol{z}), \mathfrak{U}\rangle \\
&\quad + \alpha_1\langle\mathrm{div}[\mathcal{N}(\boldsymbol{u}_1 + \boldsymbol{z}) - \mathcal{N}(\boldsymbol{u}_2 + \boldsymbol{z})], \mathfrak{U}\rangle + \alpha_2\left\langle\mathrm{div}[\mathrm{A}(\boldsymbol{u}_1 + \boldsymbol{z})^2 - \mathrm{A}(\boldsymbol{u}_2 + \boldsymbol{z})^2], \mathfrak{U}\right\rangle \\
&= -\underbrace{b(\mathfrak{U}, \boldsymbol{u}_1 + \boldsymbol{z}, \mathfrak{U})}_{=:U_1} - \underbrace{\frac{\alpha_1}{2}\langle\mathcal{N}(\boldsymbol{u}_1 + \boldsymbol{z}) - \mathcal{N}(\boldsymbol{u}_2 + \boldsymbol{z}), \mathrm{A}(\mathfrak{U})\rangle}_{=:U_2} \\
&\quad - \underbrace{\frac{\alpha_2}{2}\left(\mathrm{A}(\mathfrak{U})\mathrm{A}(\boldsymbol{u}_1 + \boldsymbol{z}) + \mathrm{A}(\boldsymbol{u}_2 + \boldsymbol{z})\mathrm{A}(\mathfrak{U}), \mathrm{A}(\mathfrak{U})\right)}_{=:U_3}.
\end{aligned} \tag{89}$$

Let us write the term U_2 in more simple way by using integration by parts as follows (see [6, Equations (31)-(32)] for detailed calculation):

$$U_2 = \frac{\alpha_1}{2}\int_{\mathbb{T}^2}[(\boldsymbol{u}_1 + \boldsymbol{z})\cdot\nabla\mathrm{A}(\boldsymbol{u}_1 + \boldsymbol{z}) - (\boldsymbol{u}_2 + \boldsymbol{z})\cdot\nabla\mathrm{A}(\boldsymbol{u}_2 + \boldsymbol{z})] : \mathrm{A}(\mathfrak{U})\mathrm{d}x$$

$$
\begin{aligned}
&+\frac{\alpha_1}{2}\int_{\mathbb{T}^2}[L(\boldsymbol{u}_1+\boldsymbol{z})^T\mathrm{A}(\boldsymbol{u}_1+\boldsymbol{z})+\mathrm{A}(\boldsymbol{u}_1+\boldsymbol{z})L(\boldsymbol{u}_1+\boldsymbol{z})\\
&-L(\boldsymbol{u}_2+\boldsymbol{z})^T\mathrm{A}(\boldsymbol{u}_2+\boldsymbol{z})+\mathrm{A}(\boldsymbol{u}_2+\boldsymbol{z})L(\boldsymbol{u}_2+\boldsymbol{z})]:\mathrm{A}(\mathfrak{U})\mathrm{d}x\\
=&\underbrace{\frac{\alpha_1}{2}\int_{\mathbb{T}^2}[\mathfrak{U}\cdot\nabla\mathrm{A}(\boldsymbol{u}_2+\boldsymbol{z})]:\mathrm{A}(\mathfrak{U})}_{=:U_{21}}\\
&\underbrace{+\alpha_1\int_{\mathbb{T}^2}[\mathrm{A}(\mathfrak{U})^2:L(\boldsymbol{u}_1+\boldsymbol{z})+(\mathrm{A}(\boldsymbol{u}_2+\boldsymbol{z})\mathrm{A}(\mathfrak{U})):L(\mathfrak{U})]}_{=:U_{22}}. \qquad (90)
\end{aligned}
$$

Using Hölder's, Sobolev's and Young's inequalities, we find

$$
|U_1|\leq\|\nabla(\boldsymbol{u}_1+\boldsymbol{z})\|_2\|\mathfrak{U}\|_4^2\leq C\|\nabla(\boldsymbol{u}_1+\boldsymbol{z})\|_2\|\nabla\mathfrak{U}\|_2^2, \qquad (91)
$$

$$
\begin{aligned}
|U_3+U_{22}|&\leq C\int_{\mathbb{T}^2}|\mathrm{A}(\mathfrak{U}(x))|(|\mathrm{A}(\boldsymbol{u}_1(x)+\boldsymbol{z}(x))|+|\mathrm{A}(\boldsymbol{u}_2(x)+\boldsymbol{z}(x))|)|\nabla\mathfrak{U}(x)|\mathrm{d}x\\
&\leq\frac{\beta}{12}\int_{\mathbb{T}^2}|\mathrm{A}(\mathfrak{U}(x))|^2(|\mathrm{A}(\boldsymbol{u}_1(x)+\boldsymbol{z}(x))|^2+|\mathrm{A}(\boldsymbol{u}_2(x)+\boldsymbol{z}(x))|^2)\mathrm{d}x\\
&\quad+C\|\nabla\mathfrak{U}\|_2^2, \qquad (92)
\end{aligned}
$$

where $\gamma>0$ will be chosen later. Finally, we estimate U_{21} which is bit more technical than other terms. Let $\sigma\in(0,1)$ be arbitrary. Then, using Hölder's inequality, (87) and Young's inequality, we get

$$
\begin{aligned}
|U_{21}|&\leq C\int_{\mathbb{T}^2}|\mathfrak{U}(x)||\nabla\mathrm{A}(\boldsymbol{u}_2(x)+\boldsymbol{z}(x))||\mathrm{A}(\mathfrak{U}(x))|\mathrm{d}x\\
&=C\int_{\mathbb{T}^2}|\mathfrak{U}(x)||\mathrm{A}(\mathfrak{U}(x))|^{\sigma}|\nabla\mathrm{A}(\boldsymbol{u}_2(x)+\boldsymbol{z}(x))||\mathrm{A}(\mathfrak{U}(x))|^{1-\sigma}\mathrm{d}x\\
&\leq C\|\mathfrak{U}\|_{\frac{4}{1-\sigma}}\|\mathrm{A}(\mathfrak{U})\|_2^{\sigma}\|\nabla\mathrm{A}(\boldsymbol{u}_2+\boldsymbol{z})\|_2\|\mathrm{A}(\mathfrak{U})\|_4^{1-\sigma}\\
&\leq\frac{C}{1-\sigma}\|\nabla\mathfrak{U}\|_2^{1+\sigma}\|\mathcal{A}(\boldsymbol{u}_2+\boldsymbol{z})\|_2\|\mathrm{A}(\mathfrak{U})\|_4^{1-\sigma}\\
&\leq\frac{\beta}{12}\|\mathrm{A}(\mathfrak{U})\|_4^4+\frac{C}{1-\sigma}\left(\frac{\beta}{3}\right)^{\frac{\sigma-1}{\sigma+3}}\|\nabla\mathfrak{U}\|_2^{\frac{4(1+\sigma)}{\sigma+3}}\|\mathcal{A}(\boldsymbol{u}_2+\boldsymbol{z})\|_2^{\frac{4}{\sigma+3}}\\
&\leq\frac{\beta}{6}\int_{\mathbb{T}^2}|\mathrm{A}(\mathfrak{U})|^2(|\mathrm{A}(\boldsymbol{u}_1+\boldsymbol{z})|^2+|\mathrm{A}(\boldsymbol{u}_2+\boldsymbol{z})|^2)\\
&\quad+\frac{C}{1-\sigma}\left(\frac{\beta}{3}\right)^{\frac{\sigma-1}{\sigma+3}}\|\nabla\mathfrak{U}\|_2^{\frac{4(1+\sigma)}{\sigma+3}}\|\mathcal{A}(\boldsymbol{u}_2+\boldsymbol{z})\|_2^{\frac{4}{\sigma+3}}. \qquad (93)
\end{aligned}
$$

Combining (89)-(93), we infer

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\mathfrak{U}(t)\|_2^2+\alpha_1\|\nabla\mathfrak{U}(t)\|_2^2\right]\\
&\leq C[1+\|\nabla(\boldsymbol{u}_1(t)+\boldsymbol{z}(t))\|_2]\|\nabla\mathfrak{U}(t)\|_2^2\\
&\quad+\frac{C}{1-\sigma}\left(\frac{\beta}{3}\right)^{\frac{\sigma-1}{\sigma+3}}\|\nabla\mathfrak{U}(t)\|_2^{\frac{4(1+\sigma)}{\sigma+3}}\|\mathcal{A}(\boldsymbol{u}_2(t)+\boldsymbol{z}(t))\|_2^{\frac{4}{\sigma+3}},\ \text{for a.e. } t\in[0,T]. \qquad (94)
\end{aligned}
$$

In view of (60), there exists a constant $\widehat{K} := \widehat{K}(\alpha_1, \alpha_2, \beta, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{z}, T) > 0$ independent of σ such that

$$\frac{\mathrm{d}\mathbf{U}(t)}{\mathrm{d}t} \leq \frac{\widehat{K}}{1-\sigma}\bigg[\mathbf{U}(t) + \{\mathbf{U}(t)\}^{\frac{2(1+\sigma)}{\sigma+3}}\bigg] \tag{95}$$

for a.e. $t \in [0,T]$, where $\mathbf{U}(t) = \|\mathfrak{U}(t)\|_2^2 + \alpha_1\|\nabla\mathfrak{U}(t)\|_2^2$. We also have

$$\frac{\mathrm{d}\mathbf{U}(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg[\{\mathbf{U}(t)\}^{\frac{1-\sigma}{\sigma+3}}\bigg]^{\frac{\sigma+3}{1-\sigma}} = \frac{\sigma+3}{1-\sigma}\{\mathbf{U}(t)\}^{\frac{2(1+\sigma)}{\sigma+3}}\frac{\mathrm{d}}{\mathrm{d}t}\bigg[\{\mathbf{U}(t)\}^{\frac{1-\sigma}{\sigma+3}}\bigg], \tag{96}$$

for a.e. $t \in [0,T]$, which implies from (95) and $\mathbf{U}(t) \geq 0$ that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg[\{\mathbf{U}(t)\}^{\frac{1-\sigma}{\sigma+3}}\bigg] \leq \frac{\widehat{K}}{\sigma+3}\bigg[\{\mathbf{U}(t)\}^{\frac{1-\sigma}{\sigma+3}} + 1\bigg], \tag{97}$$

for a.e. $t \in [0,T]$. Now, an application of Gronwall's inequality to (97) gives

$$\mathbf{U}(t) \leq \left\{\frac{\widehat{K}t}{\sigma+3}\right\}^{\frac{\sigma+3}{1-\sigma}} e^{\widehat{K}t(1-\sigma)} \leq \left\{\frac{\widehat{K}t}{3}\right\}^{\frac{\sigma+3}{1-\sigma}} e^{\widehat{K}t(1-\sigma)}, \tag{98}$$

for all $t \in [0,T]$. Let us denote $T_0 = \frac{3}{2\widehat{K}}$, then we have

$$\mathbf{U}(t) \leq \left\{\frac{1}{2}\right\}^{\frac{\sigma+3}{1-\sigma}} e^{\frac{3}{2}}, \quad \text{for all } t \in [0,T_0]. \tag{99}$$

If $\mathbf{U}(t) \geq \varepsilon$ for some $\varepsilon > 0$. Then, one can always find $\sigma \in (0,1)$ close the 1 such that

$$\mathbf{U}(t) \leq \left\{\frac{1}{2}\right\}^{\frac{\sigma+3}{1-\sigma}} e^{\frac{3}{2}} \leq \frac{\varepsilon}{2}, \quad \text{for all } t \in [0,T_0], \tag{100}$$

which leads to a contradiction. Hence $\mathbf{U}(t) = 0$ for all $t \in [0,T_0]$. We repeat the same arguments a finite number of times to cover the whole interval $[0,T]$, so that the uniqueness is proved.

3.1.6. *More regular a priori estimate.* We will denote by $\mathfrak{D}^k\boldsymbol{w}$ a linear combination of derivatives of order not greater than k of the components of $\boldsymbol{w}$. The expression $\mathfrak{D}^k\boldsymbol{w}_1\mathfrak{D}^l\boldsymbol{w}_2$ denotes a product of terms of the type $\mathfrak{D}^k\boldsymbol{w}_1$ and $\mathfrak{D}^l\boldsymbol{w}_2$ or a sum of such terms. Let us recall an inequality from [4, Proposition 1] which will be used in the sequel: let $\varepsilon_0 > 0$, then for any $\boldsymbol{v} \in \dot{\mathbb{H}}^{1+\varepsilon_0}(\mathbb{T}^2)$

$$\|\boldsymbol{v}\|_\infty \leq \frac{C_*}{\sqrt{\varepsilon_0}}\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^{1+\varepsilon_0}}, \tag{101}$$

where C_* is a universal constant independent of ε_0.

Since $\{\boldsymbol{w}_i\}_{i\in\mathbb{N}}$ is a sequence of eigenfunctions of the Stokes operator $\mathcal{A}$, it gives $\mathcal{A}^2\boldsymbol{u}_n \in \mathbb{H}_n$. Taking the inner product with $\mathcal{A}^2\boldsymbol{u}_n$ in $(37)_1$, we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\bigg[\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(t)\|_2^2\bigg] + \nu\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(t)\|_2^2 \\
&= \underbrace{-b(\boldsymbol{u}_n(t)+\boldsymbol{z}(t), \boldsymbol{u}_n(t)+\boldsymbol{z}(t), \mathcal{A}^2\boldsymbol{u}_n(t))}_{:=\widehat{I}_1} \\
&\quad \underbrace{-\alpha_1 b(\boldsymbol{u}_n(t)+\boldsymbol{z}(t), \mathcal{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)), \mathcal{A}^2(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)))}_{:=\widehat{I}_2}
\end{aligned}$$

$$
\begin{aligned}
&\underbrace{+\alpha_1 b(\boldsymbol{u}_n(t)+\boldsymbol{z}(t),\mathcal{A}(\boldsymbol{u}_n(t)+\boldsymbol{z}(t)),\mathcal{A}^2\boldsymbol{z}(t))}_{:=\widehat{I}_3}\\
&-\underbrace{\alpha_1 b(\mathcal{A}^2\boldsymbol{u}_n(t),\boldsymbol{u}_n(t)+\boldsymbol{z}(t),\mathcal{A}\boldsymbol{u}_n(t)+\mathcal{A}\boldsymbol{z}(t))}_{:=\widehat{I}_4}\\
&\underbrace{-\frac{(\alpha_1+\alpha_2)}{2}\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))]^2:\mathrm{A}(\mathcal{A}^2\boldsymbol{u}_n(x,t))\mathrm{d}x}_{:=\widehat{I}_5}\\
&\underbrace{-\frac{\beta}{2}\int_{\mathbb{T}^2}|\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t)):\mathrm{A}(\mathcal{A}^2(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t)))]\mathrm{d}x}_{:=\widehat{I}_6}\\
&\underbrace{+\frac{\beta}{2}\int_{\mathbb{T}^2}|\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t))|^2[\mathrm{A}(\boldsymbol{u}_n(x,t)+\boldsymbol{z}(x,t)):\mathrm{A}(\mathcal{A}^2\boldsymbol{z}(x,t))]\mathrm{d}x}_{:=\widehat{I}_7}\\
&+\underbrace{[(\boldsymbol{f}(t),\mathcal{A}^2\boldsymbol{u}_n(t))+(\boldsymbol{z}(t)+(\theta+\alpha_1-\nu)\mathcal{A}\boldsymbol{z}(t)+\theta\alpha_1\mathcal{A}^2\boldsymbol{z}(t),\mathcal{A}^2\boldsymbol{u}_n(t))]}_{:=\widehat{I}_8}. \qquad (102)
\end{aligned}
$$

Using integration by parts, let us rewrite each term of right hand side of (102) one by one in terms of derivatives $\mathfrak{D}^k$ as follows:

$$
\begin{aligned}
\widehat{I}_1 &= \int_{\mathbb{T}^2}\mathfrak{D}^0(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3\boldsymbol{u}_n\mathrm{d}x\\
&\quad+\int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3\boldsymbol{u}_n\mathrm{d}x, && (103)\\
\widehat{I}_2 &= \int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x, && (104)\\
\widehat{I}_3 &= \int_{\mathbb{T}^2}\mathfrak{D}^0(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^4\boldsymbol{z}\mathrm{d}x, && (105)\\
\widehat{I}_4 &= \int_{\mathbb{T}^2}\mathfrak{D}^3\boldsymbol{u}_n\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x\\
&\quad+\int_{\mathbb{T}^2}\mathfrak{D}^3\boldsymbol{u}_n\cdot\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x, && (106)\\
\widehat{I}_5 &= \int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3\boldsymbol{u}_n\mathrm{d}x, && (107)\\
\widehat{I}_7 &= \underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3\boldsymbol{z}\mathrm{d}x}_{\widehat{I}_{71}}\\
&\quad+\underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3\boldsymbol{z}\mathrm{d}x}_{\widehat{I}_{72}}, && (108)\\
\widehat{I}_6 &= -\frac{\beta}{2}\int_{\mathbb{T}^2}|\Delta\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})|^2|\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})|^2\mathrm{d}x-\beta\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z}):\Delta\mathrm{A}(\boldsymbol{u}_n+\boldsymbol{z})]^2\mathrm{d}x\\
&\quad+\underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x}_{\widehat{I}_{61}}. && (109)
\end{aligned}
$$

Equalities (102)-(109) provide that

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\mathcal{A}\boldsymbol{u}_n\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n\|_2^2\right]+\nu\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n\|_2^2\\
&\leq \widehat{I}_8+\widehat{I}_{71}+\widehat{I}_{72}+\widehat{I}_{61}+\underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x}_{\widehat{I}_9}\\
&\quad+\underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^4\boldsymbol{z}\mathrm{d}x}_{\widehat{I}_{10}}\\
&\quad+\underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x}_{\widehat{I}_{11}}\\
&\quad+\underbrace{\int_{\mathbb{T}^2}\mathfrak{D}^3\boldsymbol{z}\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\cdot\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\mathrm{d}x}_{\widehat{I}_{12}}. \qquad (110)
\end{aligned}
$$

Let us now estimate each term of right hand side of (110) using Hölder's, Sobolev's, (101), interpolation and Young's inequalities, as follows:

$$
|\widehat{I}_8|\leq C\|\boldsymbol{f}\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^5}^2+\frac{\nu}{2}\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n\|_2^2, \qquad (111)
$$

$$
\begin{aligned}
|\widehat{I}_{71}|&\leq C\|\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\|_\infty^2\|\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\mathfrak{D}^3\boldsymbol{z}\|_2\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^{2+\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^{1-\varepsilon_0}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^{2+\varepsilon_0}\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}, \qquad (112)
\end{aligned}
$$

$$
\begin{aligned}
|\widehat{I}_{72}|&\leq C\|\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\|_\infty\|\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\|_4^2\|\mathfrak{D}^3\boldsymbol{z}\|_2\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^{2+\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^{1-\varepsilon_0}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^{2+\varepsilon_0}\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}, \qquad (113)
\end{aligned}
$$

$$
\begin{aligned}
|\widehat{I}_{61}|&\leq C\|\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\|_\infty\|\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\|_4^2\|\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\|_2\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^{2+\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^{\frac{5}{2}}}^2\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^{2-\varepsilon_0}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^{2+\varepsilon_0}, \qquad (114)
\end{aligned}
$$

$$
\begin{aligned}
|\widehat{I}_9|&\leq C\|\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\|_\infty\|\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\|_2^2\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^{2+\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2\\
&\leq\frac{C}{\sqrt{\varepsilon_0}}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}^{1-\varepsilon_0}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^{2+\varepsilon_0}, \qquad (115)
\end{aligned}
$$

$$
|\widehat{I}_{10}|\leq C\|\mathfrak{D}^1(\boldsymbol{u}_n+\boldsymbol{z})\|_\infty\|\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\|_2\|\mathfrak{D}^4\boldsymbol{z}\|_2\leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^4}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2, \qquad (116)
$$

$$
|\widehat{I}_{11}|\leq C\|\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\|_4^2\|\mathfrak{D}^3(\boldsymbol{u}_n+\boldsymbol{z})\|_2
$$

$$
\begin{aligned}
&\leq C\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^{\frac{5}{2}}}^2\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\\
&\leq C\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^2}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2, \qquad (117)
\end{aligned}
$$

$$
|\widehat{I}_{12}|\leq C\|\mathfrak{D}^2(\boldsymbol{u}_n+\boldsymbol{z})\|_4^2\|\mathfrak{D}^3\boldsymbol{z}\|_2\leq C\|\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}\|\boldsymbol{u}_n+\boldsymbol{z}\|_{\dot{\mathbb{H}}^3}^2. \qquad (118)
$$

Making use of (111)-(118) in (110), we reach at

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(t)\|_2^2\right]+2\nu\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(t)\|_2^2\\
&\leq C\|\boldsymbol{f}(t)\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^2+C\left[\|\boldsymbol{u}_n(t)+\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^2}+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^4}\right]\|\boldsymbol{u}_n(t)+\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^3}^2\\
&\quad+\frac{C}{\sqrt{\varepsilon_0}}\left[\|\boldsymbol{u}_n(t)+\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^2}^{1-\varepsilon_0}\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^3}+\|\boldsymbol{u}_n(t)+\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^2}^{2-\varepsilon_0}+\|\boldsymbol{u}_n(t)+\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^2}^{1-\varepsilon_0}\right]\\
&\quad\times\|\boldsymbol{u}_n(t)+\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^3}^{2+\varepsilon_0}, \qquad (119)
\end{aligned}
$$

which implies that $\mathbf{Y}(t)=1+\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(t)\|_2^2$ satisfies

$$
\begin{aligned}
\frac{\mathrm{d}\mathbf{Y}(t)}{\mathrm{d}t}&\leq C\left[1+\|\boldsymbol{f}(t)\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^6\right]\\
&\quad+\frac{C}{\sqrt{\varepsilon_0}}\left[1+\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^6\right]\{\mathbf{Y}(t)\}^{1+\frac{\varepsilon_0}{2}}. \qquad (120)
\end{aligned}
$$

Therefore, we obtain

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\{\mathbf{Y}(t)\}^{-\frac{\varepsilon_0}{2}}\right]&=-\frac{\varepsilon_0}{2\mathbf{Y}(t)^{1+\frac{\varepsilon_0}{2}}}\frac{\mathrm{d}\mathbf{Y}(t)}{\mathrm{d}t}\\
&\geq-\frac{\varepsilon_0}{2\mathbf{Y}(t)^{1+\frac{\varepsilon_0}{2}}}\bigg[C\big[1+\|\boldsymbol{f}(t)\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^6\big]\\
&\quad+\frac{C}{\sqrt{\varepsilon_0}}\big[1+\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^6\big]\{\mathbf{Y}(t)\}^{1+\frac{\varepsilon_0}{2}}\bigg]\\
&\geq-C\sqrt{\varepsilon_0}\left[1+\|\boldsymbol{f}(t)\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}(t)\|_{\dot{\mathbb{H}}^5}^6+\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2\right], \qquad (121)
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\{\mathbf{Y}(t)\}^{-\frac{\varepsilon_0}{2}}\\
&\geq\{\mathbf{Y}(0)\}^{-\frac{\varepsilon_0}{2}}-C\sqrt{\varepsilon_0}\int_0^t\left[1+\|\boldsymbol{f}(s)\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^5}^6+\|\mathcal{A}\boldsymbol{u}_n(s)\|_2^2\right]\mathrm{d}s. \qquad (122)
\end{aligned}
$$

In view of (60) and the fact that $\boldsymbol{f}\in\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))$ and $\boldsymbol{z}\in\mathrm{C}([0,T];\dot{\mathbb{H}}^5(\mathbb{T}^2))$, we infer that, for any given time T, there exists $0<\varepsilon_*=\varepsilon_*(T,\omega)\leq 1$ such that for any $t\in[0,T]$

$$
C\int_0^t\left[1+\|\boldsymbol{f}(s)\|_{\dot{\mathbb{H}}^1}^2+\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^5}^6+\|\mathcal{A}\boldsymbol{u}_n(s)\|_2^2\right]\mathrm{d}s\leq\frac{{\varepsilon_*}^{-\frac{1}{2}}}{2}\{\mathbf{Y}(0)\}^{-\frac{\varepsilon_*}{2}}. \qquad (123)
$$

Note that such an ε_* exists as the limit when $\varepsilon_0\to 0$ of the right-hand side is $+\infty$. Moreover, one can choose ε_* small enough which will be suitable for any given $T>0$ and any given $\omega\in\widetilde{\Omega}$, that is, independent of T and ω. Hence, we obtain

$$
\{\mathbf{Y}(t)\}^{-\frac{\varepsilon_*}{2}}\geq\frac{\{\mathbf{Y}(0)\}^{-\frac{\varepsilon_*}{2}}}{2}, \qquad (124)
$$

for $t\in[0,T]$, equivalently

$$
\begin{aligned}
\|\mathcal{A}\boldsymbol{u}_n(t)\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(t)\|_2^2 &\leq 4^{\frac{1}{\varepsilon_*}}(1+\|\mathcal{A}\boldsymbol{u}_n(0)\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}_n(0)\|_2^2)\\
&\leq 4^{\frac{1}{\varepsilon_*}}(1+\|\mathcal{A}\boldsymbol{u}(0)\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}(0)\|_2^2)\\
&\leq 4^{\frac{1}{\varepsilon_*}}(1+\|\mathcal{A}\boldsymbol{v}_0\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{v}_0\|_2^2), \qquad (125)
\end{aligned}
$$

for all $t\in[0,T]$. In addition, due to (125), one can immediately obtain that

$$
\|\partial_t(\mathrm{I}+\alpha_1\mathcal{A})\boldsymbol{u}_n\|_{\mathrm{L}^2(0,T;\mathbb{H})}\leq C^*, \qquad (126)
$$

where $C^*>0$ is a constant independent of n.

This provides that $\{\boldsymbol{u}_n\}_{n\in\mathbb{N}}$ is bounded in $\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))\cap\mathrm{H}^1(0,T;\mathrm{D}(\mathcal{A}))$, consequently the unique solution $\boldsymbol{u}$ to system (17) belongs to $\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))\cap\mathrm{H}^1(0,T;\mathrm{D}(\mathcal{A}))$ satisfying

$$
\|\mathcal{A}\boldsymbol{u}(t)\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{u}(t)\|_2^2\leq 4^{\frac{1}{\varepsilon_*}}(1+\|\mathcal{A}\boldsymbol{v}_0\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{v}_0\|_2^2), \qquad (127)
$$

for all $t\in[0,T]$.

3.1.7. *Regularity of solution of system* (31). Since, $\boldsymbol{v}_0\in\mathrm{D}(\mathcal{A}^{\frac{3}{2}})$ is deterministic, we achieve from (127) that $\boldsymbol{u}\in\mathrm{L}^p(\Omega;\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}})))$, for any $p\geq 2$. Hence, in view of (18), we have $\boldsymbol{v}\in\mathrm{L}^p(\Omega;\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}})))$, $p\geq 2$. On the other hand, for given $\hat{c}>0$, let us choose $\theta>\hat{c}+\frac{\|GG^*\|_{\mathcal{O}_P}}{\lambda_1}$ and consider

$$
\begin{aligned}
&\mathbb{E}\bigg[\exp\bigg\{\hat{c}\int_0^t\|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\bigg\}\bigg]\\
&\leq\mathbb{E}\bigg[\exp\bigg\{2\hat{c}\bigg[\int_0^t\|\boldsymbol{u}(s)\|_{\dot{\mathbb{H}}^3}^2+\int_0^t\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\bigg]\mathrm{d}s\bigg\}\bigg]\\
&\leq\mathbb{E}\bigg[\exp\bigg\{2\hat{c}\int_0^t\|\boldsymbol{u}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\bigg\}\times\exp\bigg\{2\hat{c}\int_0^t\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\bigg\}\bigg]\\
&\leq\mathbb{E}\bigg[\exp\bigg\{\frac{2\hat{c}t}{\alpha_1}4^{\frac{1}{\varepsilon_*}}(1+\|\mathcal{A}\boldsymbol{v}_0\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{v}_0\|_2^2)\bigg\}\times\exp\bigg\{2\hat{c}\int_0^t\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\bigg\}\bigg]\\
&\leq\exp\bigg\{\frac{2\hat{c}t}{\alpha_1}4^{\frac{1}{\varepsilon_*}}(1+\|\mathcal{A}\boldsymbol{v}_0\|_2^2+\alpha_1\|\mathcal{A}^{\frac{3}{2}}\boldsymbol{v}_0\|_2^2)\bigg\}\mathbb{E}\bigg[\exp\bigg\{2\hat{c}\int_0^t\|\boldsymbol{z}(s)\|_{\dot{\mathbb{H}}^3}^2\mathrm{d}s\bigg\}\bigg]\\
&<+\infty, \qquad (128)
\end{aligned}
$$

for all $t\in[0,T]$, where we have used (127) and (19). This completes the proof of Theorem 3.2.

4. **Linearized state equation.** In this section, we establish well-posedness results for the linearized state equation. Let T represent the time horizon for which the unique solution $\boldsymbol{v}$ to the state equation (31), as defined in Definition 3.4, exists and satisfies (33). Additionally, recall that $\mathrm{A}(\boldsymbol{v})=\nabla\boldsymbol{v}+(\nabla\boldsymbol{v})^T$, with its components denoted as $[\mathrm{A}(\boldsymbol{v})]_{ij}:=a_{ij}$. Let $\boldsymbol{\psi}:\Omega\times\mathbb{T}^2\times[0,T]\to\mathbb{R}^2$ be a force such that $\boldsymbol{\psi}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>2$) and consider the following system:

$$
\begin{cases}
\partial_t(\mathfrak{m}-\alpha_1\Delta\mathfrak{m}) = \boldsymbol{\psi} - \nabla\pi + \nu\Delta\mathfrak{m} - (\boldsymbol{v}\cdot\nabla)(\mathfrak{m}-\alpha_1\Delta\mathfrak{m}) - (\mathfrak{m}\cdot\nabla)(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v}) \\
\qquad - \sum_{j=1}^{2}[\mathfrak{m}-\alpha_1\Delta\mathfrak{m}]^j\nabla\boldsymbol{v}^j \quad - \sum_{j=1}^{2}[\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v}]^j\nabla\mathfrak{m}^j \\
\qquad + (\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{m}) + \mathrm{A}(\mathfrak{m})\mathrm{A}(\boldsymbol{v})) \\
\qquad + \beta[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{m}))] + 2\beta\mathrm{div}[(\mathrm{A}(\mathfrak{m}):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})], & \text{in } \mathbb{T}^2\times(0,\infty), \\
\mathrm{div}\ \mathfrak{m} = 0, & \text{in } \mathbb{T}^2\times[0,\infty), \\
\mathfrak{m}(x,0) = \mathbf{0}, & \text{in } \mathbb{T}^2,
\end{cases}
\tag{129}
$$

with periodic boundary conditions, where $\mathfrak{m}=(\mathfrak{m}^1,\mathfrak{m}^2)$ and π are unknown vector and scalar fields, respectively. Applying the projection $\mathcal{P}$ to the system (129), we obtain

$$
\begin{cases}
\partial_t\Upsilon(\mathfrak{m}) = \mathcal{P}\boldsymbol{\psi} - \nu\mathcal{A}\mathfrak{m} - \mathcal{P}[(\boldsymbol{v}\cdot\nabla)\Upsilon(\mathfrak{m})] - \mathcal{P}[(\mathfrak{m}\cdot\nabla)\Upsilon(\boldsymbol{v})] \\
\qquad - \mathcal{P}\Big[\sum_{j=1}^{2}[\Upsilon(\mathfrak{m})]^j\nabla\boldsymbol{v}^j\Big] - \mathcal{P}\Big[\sum_{j=1}^{2}[\Upsilon(\boldsymbol{v})]^j\nabla\mathfrak{m}^j\Big] \\
\qquad + (\alpha_1+\alpha_2)\mathcal{P}[\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{m}) + \mathrm{A}(\mathfrak{m})\mathrm{A}(\boldsymbol{v}))] \\
\qquad + \beta\mathcal{P}[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{m}))] + 2\beta\mathcal{P}[\mathrm{div}[(\mathrm{A}(\mathfrak{m}):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})]], \\
\mathfrak{m}(0) = \mathbf{0},
\end{cases}
\tag{130}
$$

in $\mathbb{V}'$.

Definition 4.1. Assume that $\boldsymbol{\psi}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>2$). A stochastic process $\{\mathfrak{m}(t)\}_{t\in[0,T]}$ with trajectories in $\mathrm{C}([0,T];\mathbb{V})\cap\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$, $\partial_t\mathfrak{m}\in\mathrm{L}^2(0,T;\mathbb{V})$ and $\mathfrak{m}(0)=\mathbf{0}$, is called a *solution* to system (130) if for any $\phi\in\mathbb{V}$ the following equation holds

$$
\begin{aligned}
&\langle\Upsilon(\mathfrak{m}(t)),\phi\rangle \\
&= -\int_0^t\langle\nu\mathcal{A}\mathfrak{m}(s) - (\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v}(s))\mathrm{A}(\mathfrak{m}(s)) + \mathrm{A}(\mathfrak{m}(s))\mathrm{A}(\boldsymbol{v}(s))) \\
&\quad - \beta\mathrm{div}(|\mathrm{A}(\boldsymbol{v}(s))|^2\mathrm{A}(\mathfrak{m}(s))) - 2\beta\mathrm{div}((\mathrm{A}(\mathfrak{m}(s)):\mathrm{A}(\boldsymbol{v}(s)))\mathrm{A}(\boldsymbol{v}(s))),\phi(s)\rangle\mathrm{d}s \\
&\quad + \int_0^t b(\boldsymbol{v}(s),\phi(s),\Upsilon(\mathfrak{m}(s)))\mathrm{d}s + \int_0^t b(\mathfrak{m}(s),\phi(s),\Upsilon(\boldsymbol{v}(s)))\mathrm{d}s \\
&\quad - \int_0^t b(\phi(s),\boldsymbol{v}(s),\Upsilon(\mathfrak{m}(s)))\mathrm{d}s - \int_0^t b(\phi(s),\mathfrak{m}(s),\Upsilon(\boldsymbol{v}(s)))\mathrm{d}s,
\end{aligned}
\tag{131}
$$

for all $t\in[0,T]$ and $\mathbb{P}$-a.s.

Theorem 4.2. *Let $T>0$, $\boldsymbol{\psi}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>2$) and $\boldsymbol{v}$ be the solution to the state equation (31) in the sense of Definition 3.1 satisfying (33). Then, there exists a unique solution $\mathfrak{m}\in\mathrm{L}^2(\Omega;\mathrm{L}^\infty(0,T;\mathbb{V}))$ to system (130) in the sense of Definition 4.1.*

Proof of Theorem 4.2. Due to assumptions, we have $\boldsymbol{\psi}\in\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))$ and $\boldsymbol{v}\in\mathrm{L}^\infty(0,T;\dot{\mathbb{H}}^3(\mathbb{T}^2))$, $\mathbb{P}$-a.s. Therefore, we will consider the system (130) pathwise and do the analysis accordingly. After showing the pathwise well-posedness, we will verify that the second-order moments are finite. The proof is divided into the following five steps.

Step I. *Finite-dimensional approximation.* Let us consider the following approximate equation for system (130) on the finite-dimensional space $\mathbb{H}_n$:

$$\begin{cases} \partial_t(\Upsilon(\mathfrak{m}_n)) = \mathrm{P}_n\mathcal{P}\boldsymbol{\psi} - \nu\mathcal{A}\mathfrak{m}_n - \mathrm{P}_n\mathcal{P}[(\boldsymbol{v}\cdot\nabla)\Upsilon(\mathfrak{m}_n)] - \mathcal{P}[(\mathfrak{m}_n\cdot\nabla)\Upsilon(\boldsymbol{v})] \\ \qquad\qquad - \mathrm{P}_n\mathcal{P}\left[\sum_{j=1}^{2}[\Upsilon(\mathfrak{m}_n)]^j\nabla\boldsymbol{v}^j\right] - \mathrm{P}_n\mathcal{P}\left[\sum_{j=1}^{2}[\Upsilon(\boldsymbol{v})]^j\nabla(\mathfrak{m}_n)^j\right] \\ \qquad\qquad + (\alpha_1+\alpha_2)\mathrm{P}_n\mathcal{P}[\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{m}_n)+\mathrm{A}(\mathfrak{m}_n)\mathrm{A}(\boldsymbol{v}))] \\ \qquad\qquad + \beta\mathrm{P}_n\mathcal{P}[[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{m}_n))]] + 2\beta\mathrm{P}_n\mathcal{P}[\mathrm{div}[(\mathrm{A}(\mathfrak{m}_n):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})]], \\ \mathfrak{m}_n(0) = \mathbf{0}, \end{cases} \tag{132}$$

for a.e. $t\in[0,T]$. Note that system (132) defines a system of linear ordinary differential equations, which has a unique local solution $\mathfrak{m}_n\in\mathrm{C}([0,T_n];\mathbb{H}_n)$, for some $0<T_n\leq T$. The following uniform a priori estimates show that the time $T_n=T$.

Step II. *A priori estimates.* Taking the inner product with $\mathfrak{m}_n$ to $(132)_1$, integrating by parts and using (11), we obtain

$$\begin{aligned} &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\mathfrak{m}_n(t)\|_2^2 + \alpha_1\|\nabla\mathfrak{m}_n(t)\|_2^2\right] + \nu\|\nabla\mathfrak{m}_n(t)\|_2^2 + \frac{\beta}{2}\||\mathrm{A}(\boldsymbol{v}(t))|\mathrm{A}(\mathfrak{m}_n(t))\|_2^2 \\ &+ \beta\int_{\mathbb{T}^2}(\mathrm{A}(\mathfrak{m}_n(x,t)):\mathrm{A}(\boldsymbol{v}(x,t)))^2\mathrm{d}x \\ &= \underbrace{(\boldsymbol{\psi}(t),\mathfrak{m}_n(t))}_{=:M_1} - \underbrace{\alpha_1 b(\boldsymbol{v}(t),\mathfrak{m}_n(t),\Delta\mathfrak{m}_n(t))}_{=:M_2} - \underbrace{b(\mathfrak{m}_n(t),\boldsymbol{v}(t),\mathfrak{m}_n(t))}_{=:M_3} \\ &\quad + \underbrace{\alpha_1 b(\mathfrak{m}_n(t),\boldsymbol{v}(t),\Delta\mathfrak{m}_n(t))}_{=:M_4} \\ &\quad - \underbrace{(\alpha_1+\alpha_2)\left\langle\mathrm{A}(\boldsymbol{v}(t))\mathrm{A}(\mathfrak{m}_n(t)) + \mathrm{A}(\mathfrak{m}_n(t))\mathrm{A}(\boldsymbol{v}(t))),\nabla\mathfrak{m}_n(t)\right\rangle}_{=:M_5}, \end{aligned} \tag{133}$$

for a.e. $t\in[0,T]$. Using integration by parts, Hölder's, Sobolev's, Ladyzhenskaya's, Young's inequalities, we deduce

$$|M_1| \leq C\|\boldsymbol{\psi}\|_2^2 + C\|\mathfrak{m}_n\|_2^2, \tag{134}$$

$$\begin{aligned} |M_2| &= \alpha_1\left|\int_{\mathbb{T}^2}\mathfrak{D}^1\boldsymbol{v}\cdot\mathfrak{D}^1\mathfrak{m}_n\cdot\mathfrak{D}^1\mathfrak{m}_n\mathrm{d}x\right| \\ &\leq C\|\nabla\boldsymbol{v}\|_\infty\|\nabla\mathfrak{m}_n\|_2^2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\nabla\mathfrak{m}_n\|_2^2, \end{aligned} \tag{135}$$

$$|M_3| \leq C\|\nabla\boldsymbol{v}\|_\infty\|\mathfrak{m}_n\|_2^2 \leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\nabla\mathfrak{m}_n\|_2^2, \tag{136}$$

$$\begin{aligned} |M_4| &= \left|\int_{\mathbb{T}^2}\mathfrak{D}^1\mathfrak{m}_n\cdot\mathfrak{D}^1\boldsymbol{v}\cdot\mathfrak{D}^1\mathfrak{m}_n\mathrm{d}x + \int_{\mathbb{T}^2}\mathfrak{D}^0\mathfrak{m}_n\cdot\mathfrak{D}^2\boldsymbol{v}\cdot\mathfrak{D}^1\mathfrak{m}_n\right| \\ &\leq C\|\nabla\boldsymbol{v}\|_\infty\|\nabla\mathfrak{m}_n\|_2^2 + C\|\mathfrak{D}^2\boldsymbol{v}\|_4\|\mathfrak{m}_n\|_4\|\nabla\mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\nabla\mathfrak{m}_n\|_2^2, \end{aligned} \tag{137}$$

$$|M_5| \leq \frac{\beta}{4}\||\mathrm{A}(\boldsymbol{v})|\mathrm{A}(\mathfrak{m}_n)\|_2^2 + C\|\nabla\mathfrak{m}_n\|_2^2. \tag{138}$$

Combining (133)-(138), we achieve

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\bigg[\|\mathfrak{m}_n(t)\|_2^2+\alpha_1\|\nabla\mathfrak{m}_n(t)\|_2^2\bigg]\\&\leq C\|\boldsymbol{\psi}(t)\|_2^2+C\{1+\|\boldsymbol{v}(t)\|_{\dot{\mathbb{H}}^3}\}\bigg[\|\mathfrak{m}_n(t)\|_2^2+\alpha_1\|\nabla\mathfrak{m}_n(t)\|_2^2\bigg],\end{aligned}\tag{139}$$

for a.e. $t\in[0,T]$, and the Gronwall inequality implies

$$\|\mathfrak{m}_n(t)\|_2^2+\alpha_1\|\nabla\mathfrak{m}_n(t)\|_2^2\leq Ce^{Ct+C\int_0^t\|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}\mathrm{d}s}\int_0^t\|\boldsymbol{\psi}(s)\|_2^2\mathrm{d}s,\tag{140}$$

for all $t\in[0,T]$. Using the fact that $\boldsymbol{\psi}\in\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))$ and $\boldsymbol{v}\in\mathrm{L}^\infty(0,T;\dot{\mathbb{H}}^3(\mathbb{T}^2))$, we have from (140) that

$$\{\mathfrak{m}_n\}_{n\in\mathbb{N}}\text{ is a bounded sequence in }\mathrm{L}^\infty(0,T;\mathbb{V}).\tag{141}$$

Next, since $\{\boldsymbol{w}_i\}_{i\in\mathbb{N}}$ is a sequence of eigenfunctions of the Stokes operator $\mathcal{A}$, it gives $\mathcal{A}\mathfrak{m}_n\in\mathbb{H}_n$, therefore taking the inner product with $\mathcal{A}\mathfrak{m}_n$ to $(132)_1$ and using (11), we obtain

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\bigg[\|\nabla\mathfrak{m}_n(t)\|_2^2+\alpha_1\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2\bigg]+\nu\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2\\&=\underbrace{(\boldsymbol{\psi}(t),\mathcal{A}\mathfrak{m}_n(t))}_{\widetilde{M}_1}-\underbrace{b(\boldsymbol{v}(t),\mathfrak{m}_n(t),\mathcal{A}\mathfrak{m}_n(t))}_{\widetilde{M}_2}-\underbrace{b(\mathfrak{m}_n(t),\Upsilon(\boldsymbol{v}(t)),\mathcal{A}\mathfrak{m}_n(t))}_{\widetilde{M}_3}\\&\quad-\underbrace{b(\mathcal{A}\mathfrak{m}_n(t),\boldsymbol{v}(t),\Upsilon(\mathfrak{m}_n(t)))}_{\widetilde{M}_4}+\underbrace{b(\mathcal{A}\mathfrak{m}_n(t),\Upsilon(\boldsymbol{v}(t)),\mathfrak{m}_n(t))}_{\widetilde{M}_5}\\&\quad+\underbrace{(\alpha_1+\alpha_2)\int_{\mathbb{T}^2}\mathrm{div}[\mathrm{A}(\boldsymbol{v}(t))\mathrm{A}(\mathfrak{m}_n(t))+\mathrm{A}(\mathfrak{m}_n(t))\mathrm{A}(\boldsymbol{v}(t))]\cdot\mathcal{A}\mathfrak{m}_n(t)\mathrm{d}x}_{\widetilde{M}_6}\\&\quad+\underbrace{\beta\bigg[\int_{\mathbb{T}^2}\mathrm{div}[|\mathrm{A}(\boldsymbol{v}(t))|^2\mathrm{A}(\mathfrak{m}_n(t))]\cdot\mathcal{A}\mathfrak{m}_n(t)\mathrm{d}x+2\int_{\mathbb{T}^2}\mathrm{div}[(\mathrm{A}(\mathfrak{m}_n(t)):\mathrm{A}(\boldsymbol{v}(t)))\mathrm{A}(\boldsymbol{v}(t))]\cdot\mathcal{A}\mathfrak{m}_n(t)\mathrm{d}x\bigg]}_{\widetilde{M}_7},\end{aligned}\tag{142}$$

for a.e. $t\in[0,T]$. Let us estimate each term of right hand side of (142) as follows:

$$|\widetilde{M}_1|\leq C\|\boldsymbol{\psi}\|_2^2+C\|\mathcal{A}\mathfrak{m}_n\|_2^2,\tag{143}$$

$$\begin{aligned}|\widetilde{M}_2|&=\bigg|\int_{\mathbb{T}^2}\mathfrak{D}^0\boldsymbol{v}\cdot\mathfrak{D}^1\mathfrak{m}_n\cdot\mathfrak{D}^2\mathfrak{m}_n\mathrm{d}x\bigg|\\&\leq C\|\boldsymbol{v}\|_4\|\mathfrak{D}^1\mathfrak{m}_n\|_4\|\mathfrak{D}^2\mathfrak{m}_n\|_2\\&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\mathfrak{m}_n\|_2^2,\end{aligned}\tag{144}$$

$$\begin{aligned}|\widetilde{M}_3|&=\bigg|\int_{\mathbb{T}^2}\mathfrak{D}^0\mathfrak{m}_n\cdot\mathfrak{D}^3\boldsymbol{v}\cdot\mathfrak{D}^2\mathfrak{m}_n\mathrm{d}x\bigg|\\&\leq C\|\mathfrak{m}_n\|_\infty\|\mathfrak{D}^3\boldsymbol{v}\|_2\|\mathfrak{D}^2\mathfrak{m}_n\|_2\\&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\mathfrak{m}_n\|_2^2,\end{aligned}\tag{145}$$

$$\begin{aligned}|\widetilde{M}_4|&=\bigg|\int_{\mathbb{T}^2}\mathfrak{D}^2\mathfrak{m}_n\cdot\mathfrak{D}^1\boldsymbol{v}\cdot\mathfrak{D}^2\mathfrak{m}_n\mathrm{d}x\bigg|\\&\leq C\|\nabla\boldsymbol{v}\|_\infty\|\mathfrak{D}^2\mathfrak{m}_n\|_2^2\end{aligned}$$

$$\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\mathfrak{m}_n\|_2^2, \tag{146}$$

$$\begin{aligned}
|\widetilde{M}_5| &= \left|\int_{\mathbb{T}^2} \mathfrak{D}^2\mathfrak{m}_n \cdot \mathfrak{D}^3\boldsymbol{v} \cdot \mathfrak{D}^0\mathfrak{m}_n \mathrm{d}x\right| \\
&\leq C\|\mathfrak{D}^2\mathfrak{m}_n\|_2\|\mathfrak{D}^3\boldsymbol{v}\|_2\|\mathfrak{m}_n\|_\infty \\
&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\mathfrak{m}_n\|_2^2,
\end{aligned} \tag{147}$$

$$\begin{aligned}
|\widetilde{M}_6| &= \left|\int_{\mathbb{T}^2} \mathfrak{D}^1\boldsymbol{v} \cdot \mathfrak{D}^2\mathfrak{m}_n \cdot \mathfrak{D}^2\mathfrak{m}_n \mathrm{d}x + \int_{\mathbb{T}^2} \mathfrak{D}^2\boldsymbol{v} \cdot \mathfrak{D}^1\mathfrak{m}_n \cdot \mathfrak{D}^2\mathfrak{m}_n \mathrm{d}x\right| \\
&\leq C\|\mathfrak{D}^1\boldsymbol{v}\|_\infty\|\mathfrak{D}^2\mathfrak{m}_n\|_2^2 + C\|\mathfrak{D}^2\boldsymbol{v}\|_4\|\mathfrak{D}^1\mathfrak{m}_n\|_4\|\mathfrak{D}^2\mathfrak{m}_n\|_2 \\
&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}\|\mathcal{A}\mathfrak{m}_n\|_2^2,
\end{aligned} \tag{148}$$

$$\begin{aligned}
|\widetilde{M}_7| &= \left|\int_{\mathbb{T}^2} \mathfrak{D}^1\boldsymbol{v} \cdot \mathfrak{D}^2\boldsymbol{v} \cdot \mathfrak{D}^1\mathfrak{m}_n \cdot \mathfrak{D}^2\mathfrak{m}_n \mathrm{d}x + \int_{\mathbb{T}^2} \mathfrak{D}^1\boldsymbol{v} \cdot \mathfrak{D}^1\boldsymbol{v} \cdot \mathfrak{D}^2\mathfrak{m}_n \cdot \mathfrak{D}^2\mathfrak{m}_n \mathrm{d}x\right| \\
&\leq C\|\mathfrak{D}^1\boldsymbol{v}\|_\infty\|\mathfrak{D}^2\boldsymbol{v}\|_4\|\mathfrak{D}^1\mathfrak{m}_n\|_4\|\mathfrak{D}^2\mathfrak{m}_n\|_2 + C\|\mathfrak{D}^1\boldsymbol{v}\|_\infty^2\|\mathfrak{D}^2\mathfrak{m}_n\|_2^2 \\
&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2\|\mathcal{A}\mathfrak{m}_n\|_2^2.
\end{aligned} \tag{149}$$

Combining (143)-(149), we achieve

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\nabla\mathfrak{m}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2\right] \\
&\leq C\|\boldsymbol{\psi}(t)\|_2^2 + C\{1 + \|\boldsymbol{v}(t)\|_{\dot{\mathbb{H}}^3}^2\}\left[\|\nabla\mathfrak{m}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2\right],
\end{aligned} \tag{150}$$

for a.e. $t \in [0, T]$, and the Gronwall inequality implies

$$\|\nabla\mathfrak{m}_n(t)\|_2^2 + \alpha_1\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2 \leq Ce^{Ct + C\int_0^t \|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^2 \mathrm{d}s} \int_0^t \|\boldsymbol{\psi}(s)\|_2^2 \mathrm{d}s, \tag{151}$$

for all $t \in [0, T]$. Using the fact that $\boldsymbol{\psi} \in \mathrm{L}^2(0, T; \dot{\mathbb{L}}^2(\mathbb{T}^2))$ and $\boldsymbol{v} \in \mathrm{L}^\infty(0, T; \dot{\mathbb{H}}^3(\mathbb{T}^2))$, we have from (151) that

$$\{\mathfrak{m}_n\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A})). \tag{152}$$

Taking the inner product with $\partial_t\mathfrak{m}_n$ to $(132)_1$ and using integration by parts, we obtain

$$\begin{aligned}
&\|\partial_t\mathfrak{m}_n(t)\|_2^2 + \alpha_1\|\partial_t\nabla\mathfrak{m}_n(t)\|_2^2 + \frac{\nu}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla\mathfrak{m}_n(t)\|_2^2 \\
&= \underbrace{(\boldsymbol{\psi}(t), \partial_t\mathfrak{m}_n(t))}_{\widehat{M}_1} + \underbrace{b(\boldsymbol{v}(t), \partial_t\mathfrak{m}_n(t), \Upsilon(\mathfrak{m}_n(t)))}_{\widehat{M}_2} - \underbrace{b(\mathfrak{m}_n(t), \Upsilon(\boldsymbol{v}(t)), \partial_t\mathfrak{m}_n(t))}_{\widehat{M}_{31}} \\
&\quad - \underbrace{b(\partial_t\mathfrak{m}_n(t), \boldsymbol{v}(t), \Upsilon(\mathfrak{m}_n(t)))}_{\widehat{M}_4} + \underbrace{b(\partial_t\mathfrak{m}_n(t), \Upsilon(\boldsymbol{v}(t)), \mathfrak{m}_n(t))}_{\widehat{M}_{32}} \\
&\quad - \underbrace{(\alpha_1 + \alpha_2)\big(\mathrm{A}(\boldsymbol{v}(t))\mathrm{A}(\mathfrak{m}_n(t)) + \mathrm{A}(\mathfrak{m}_n(t))\mathrm{A}(\boldsymbol{v}(t)), \partial_t\nabla\mathfrak{m}_n(t)\big)}_{\widehat{M}_5} \\
&\quad - \underbrace{\beta\big(|\mathrm{A}(\boldsymbol{v}(t))|^2\mathrm{A}(\mathfrak{m}_n(t)), \partial_t\nabla\mathfrak{m}_n(t)\big)}_{\widehat{M}_6} \\
&\quad - \underbrace{2\beta\big((\mathrm{A}(\mathfrak{m}_n(t)) : \mathrm{A}(\boldsymbol{v}(t)))(\mathrm{A}(\boldsymbol{v}(t)), \partial_t\nabla\mathfrak{m}_n(t))\big)}_{\widehat{M}_7},
\end{aligned} \tag{153}$$

for a.e. $t \in [0,T]$. Let us estimate the terms on the right hand side of (153) using Hölder's, Sobolev's and Young's inequalities as follows:

$$|\widehat{M}_1| \leq C\|\boldsymbol{\psi}\|_2^2 + \frac{1}{6}\|\partial_t \mathfrak{m}\|_2^2, \tag{154}$$

$$\begin{aligned}|\widehat{M}_2| &\leq C\|\boldsymbol{v}\|_\infty \|\partial_t \nabla \mathfrak{m}_n\|_2 \|\mathfrak{D}^2 \mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\partial_t \nabla \mathfrak{m}_n\|_2 \|\mathcal{A}\mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2 \|\mathcal{A}\mathfrak{m}_n\|_2^2 + \frac{1}{6}\|\partial_t \nabla \mathfrak{m}_n\|_2^2,\end{aligned} \tag{155}$$

$$\begin{aligned}|\widehat{M}_{31} + \widehat{M}_{32}| &\leq C\|\mathfrak{D}^3 \boldsymbol{v}\|_2 \|\partial_t \mathfrak{m}_n\|_2 \|\mathfrak{m}_n\|_\infty \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\partial_t \mathfrak{m}_n\|_2 \|\mathcal{A}\mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2 \|\mathcal{A}\mathfrak{m}_n\|_2^2 + \frac{1}{6}\|\partial_t \mathfrak{m}_n\|_2^2,\end{aligned} \tag{156}$$

$$\begin{aligned}|\widehat{M}_4| &\leq C\|\nabla \boldsymbol{v}\|_\infty \|\partial_t \mathfrak{m}_n\|_2 \|\mathfrak{D}^2 \mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\partial_t \mathfrak{m}_n\|_2 \|\mathcal{A}\mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2 \|\mathcal{A}\mathfrak{m}_n\|_2^2 + \frac{1}{6}\|\partial_t \mathfrak{m}_n\|_2^2,\end{aligned} \tag{157}$$

$$\begin{aligned}|\widehat{M}_5| &\leq C\|\nabla \boldsymbol{v}\|_\infty \|\partial_t \nabla \mathfrak{m}_n\|_2 \|\nabla \mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\partial_t \nabla \mathfrak{m}_n\|_2 \|\mathcal{A}\mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2 \|\mathcal{A}\mathfrak{m}_n\|_2^2 + \frac{1}{6}\|\partial_t \nabla \mathfrak{m}_n\|_2^2,\end{aligned} \tag{158}$$

$$\begin{aligned}|\widehat{M}_6 + \widehat{M}_7| &\leq C\|\nabla \boldsymbol{v}\|_\infty^2 \|\partial_t \nabla \mathfrak{m}_n\|_2 \|\nabla \mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2 \|\partial_t \nabla \mathfrak{m}_n\|_2 \|\mathcal{A}\mathfrak{m}_n\|_2 \\ &\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^4 \|\mathcal{A}\mathfrak{m}_n\|_2^2 + \frac{1}{6}\|\partial_t \nabla \mathfrak{m}_n\|_2^2.\end{aligned} \tag{159}$$

Combining (153)-(159), we have

$$\begin{aligned}&\|\partial_t \mathfrak{m}_n(t)\|_2^2 + \alpha_1 \|\partial_t \nabla \mathfrak{m}_n(t)\|_2^2 + \nu \frac{\mathrm{d}}{\mathrm{d}t}\|\nabla \mathfrak{m}_n(t)\|_2^2 \\ &\leq C\|\boldsymbol{\psi}(t)\|_2^2 + C[1 + \|\boldsymbol{v}(t)\|_{\dot{\mathbb{H}}^3}^4]\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2,\end{aligned} \tag{160}$$

for a.e. $t \in [0,T]$, which due to $\mathfrak{m}_n(0) = \mathbf{0}$ implies

$$\begin{aligned}&\int_0^t \left[\|\partial_t \mathfrak{m}_n(s)\|_2^2 + \alpha_1 \|\partial_t \nabla \mathfrak{m}_n(s)\|_2^2\right] \mathrm{d}s \\ &\leq C \int_0^t \|\boldsymbol{\psi}(s)\|_2^2 + [1 + \|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^4]\|\mathcal{A}\mathfrak{m}_n(s)\|_2^2 \mathrm{d}s,\end{aligned} \tag{161}$$

for all $t \in [0,T]$. Using the fact that $\boldsymbol{\psi} \in \mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))$, $\boldsymbol{v} \in \mathrm{L}^\infty(0,T;\dot{\mathbb{H}}^3(\mathbb{T}^2))$ and $\{\mathfrak{m}_n\}_{n\in\mathbb{N}}$ is a bounded sequence in $\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$ (see (152)), we have from (161) that

$$\{\partial_t \mathfrak{m}_n\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^2(0,T;\mathbb{V}). \tag{162}$$

Note that, similar to (73), in view of (162), we get

$$\{\partial_t \Upsilon(\mathfrak{m}_n)\}_{n\in\mathbb{N}} \text{ is a bounded sequence in } \mathrm{L}^2(0,T;\mathbb{V}'). \tag{163}$$

Step III. *Weak and strong limits, and passing $n \to \infty$ in system* (132). Using (152), (163), the *Banach-Alaoglu theorem* and *Aubin-Lions compactness lemma* (similar to

(76)), we infer the existence of an element $\mathfrak{m}\in\mathrm{L}^{\infty}(0,T;\mathrm{D}(\mathcal{A}))$ with $\partial_t(\mathrm{I}+\alpha_1\mathcal{A})\mathfrak{m}\in\mathrm{L}^2(0,T;\mathbb{V}')$ such that

$$\left.\begin{aligned} \mathfrak{m}_n &\xrightarrow{w^*} \mathfrak{m} &&\text{in} && \mathrm{L}^{\infty}(0,T;\mathrm{D}(\mathcal{A})),\\ \partial_t\Upsilon(\mathfrak{m}_n) &\xrightarrow{w} \partial_t\Upsilon(\mathfrak{m}) &&\text{in} && \mathrm{L}^{2}(0,T;\mathbb{V}'),\\ \mathfrak{m}_n &\to \mathfrak{m} &&\text{in} && \mathrm{L}^{2}(0,T;\mathbb{V}),\end{aligned}\right\} \tag{164}$$

along a subsequence (still denoted by the same symbol). Making use of weak and strong convergence (164), $\boldsymbol{v}\in\mathrm{L}^{\infty}(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))$ and strong convergence $\mathrm{P}_n\mathcal{P}\boldsymbol{\psi}\to\mathcal{P}\boldsymbol{\psi}$ in $\mathrm{L}^2(0,T;\mathbb{H})$), one can easily pass the limit $n\to\infty$ in (132) and obtain

$$\begin{aligned}\partial_t\Upsilon(\mathfrak{m}) &= \mathcal{P}\boldsymbol{\psi}-\nu\mathcal{A}\mathfrak{m}-\mathcal{P}[(\boldsymbol{v}\cdot\nabla)\Upsilon(\mathfrak{m})]-\mathcal{P}[(\mathfrak{m}\cdot\nabla)\Upsilon(\boldsymbol{v})]-\mathcal{P}\bigg[\sum_{j=1}^{2}[\Upsilon(\mathfrak{m})]^j\nabla\boldsymbol{v}^j\bigg]\\ &\quad-\mathcal{P}\bigg[\sum_{j=1}^{2}[\Upsilon(\mathfrak{m})]^j\nabla(\mathfrak{m})^j\bigg]+(\alpha_1+\alpha_2)\mathcal{P}[\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{m})+\mathrm{A}(\mathfrak{m})\mathrm{A}(\boldsymbol{v}))]\\ &\quad+\beta\mathcal{P}[[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{m}))]]+2\beta\mathcal{P}[\mathrm{div}[(\mathrm{A}(\mathfrak{m}):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})]],\end{aligned}\tag{165}$$

in $\mathrm{L}^2(0,T;\mathbb{V}')$. Next, since $(\mathrm{I}+\alpha_1\mathcal{A})^{\frac12}\mathfrak{m}\in\mathrm{L}^2(0,T;\mathbb{V})$ and $\partial_t(\mathrm{I}+\alpha_1\mathcal{A})^{\frac12}\mathfrak{m}\in\mathrm{L}^2(0,T;\mathbb{V}')$, it implies from [31, Lemma 1.2, p.176] that $(\mathrm{I}+\alpha_1\mathcal{A})^{\frac12}\mathfrak{m}\in\mathrm{C}([0,T];\mathbb{H})$, the real-valued function $t\mapsto\|(\mathrm{I}+\alpha_1\mathcal{A})^{\frac12}\mathfrak{m}(t)\|_2^2$ is absolutely continuous and the following equality is satisfied:

$$\frac12\frac{\mathrm{d}}{\mathrm{d}t}\|(\mathrm{I}+\alpha_1\mathcal{A})^{\frac12}\mathfrak{m}(t)\|_2^2=\Big\langle\partial_t(\mathrm{I}+\alpha\mathcal{A})^{\frac12}\mathfrak{m}(t),(\mathrm{I}+\alpha_1\mathcal{A})^{\frac12}\mathfrak{m}(t)\Big\rangle,\tag{166}$$

for a.e. $t\in[0,T]$. Also, the initial condition in the Definition 4.1 is well defined.

Step IV. *Uniqueness:* Define $\mathfrak{M}:=\mathfrak{m}_1-\mathfrak{m}_2$, where $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are two solutions of system (130) in the sense of Definition 4.1. Since, system (130) is linear in $\mathfrak{m}$, therefore $\mathfrak{M}$ is also a solution of system (130) replacing $\boldsymbol{\psi}$ with $\mathbf{0}$ and it satisfies

$$\begin{aligned}&\|\mathfrak{M}(t)\|_2^2+\alpha_1\|\nabla\mathfrak{M}(t)\|_2^2+2\nu\int_0^t\|\nabla\mathfrak{M}(s)\|_2^2\mathrm{d}s+\beta\int_0^t\||\mathrm{A}(\boldsymbol{v}(s))|\mathrm{A}(\mathfrak{M}(s))\|_2^2\mathrm{d}s\\ &+2\beta\int_0^t\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{v}(s)):\mathrm{A}(\mathfrak{M}(s))]^2\,\mathrm{d}x\mathrm{d}s\\ &=-2\int_0^t[\alpha_1 b(\boldsymbol{v}(s),\mathfrak{M}(s),\Delta\mathfrak{M}(s))+b(\mathfrak{M}(s),\boldsymbol{v}(s),\mathfrak{M}(s))-\alpha_1 b(\mathfrak{M}(s),\boldsymbol{v}(s),\Delta\mathfrak{M}(s))]\mathrm{d}s\\ &\quad+(\alpha_1+\alpha_2)\int_0^t\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{v}(s))\mathrm{A}(\mathfrak{M}(s))+\mathrm{A}(\mathfrak{M}(s))\mathrm{A}(\boldsymbol{v}(s))]:\mathrm{A}(\mathfrak{M}(s))\mathrm{d}x\mathrm{d}s,\end{aligned}\tag{167}$$

for all $t\in[0,T]$. A reasoning similar to that used to obtain (139) and $\mathfrak{M}(0)=\mathbf{0}$ yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\|\mathfrak{M}(t)\|_2^2+\alpha_1\|\nabla\mathfrak{M}(t)\|_2^2\Big]\leq C\{1+\|\boldsymbol{v}(t)\|_{\dot{\mathbb{H}}^3}\}\Big[\|\mathfrak{M}(t)\|_2^2+\alpha_1\|\nabla\mathfrak{M}(t)\|_2^2\Big],\tag{168}$$

for all $t \in [0,T]$. Hence, an application of Gronwall's inequality and the fact that $\boldsymbol{v} \in \mathrm{L}^{\infty}(0,T;\mathrm{D}(\mathcal{A}^{\frac{3}{2}}))$ give $\mathfrak{m}_1(t) = \mathfrak{m}_2(t)$ in $\mathbb{V}$, for all $t \in [0,T]$.

Step V. *Moments:* Finally we show that $\mathfrak{m} \in \mathrm{L}^2(\Omega;\mathrm{L}^{\infty}(0,T;\mathrm{D}(\mathcal{A})))$. Taking supremum over $t \in [0,T]$ and then expectation of (151), we find for $p > 2$

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\|\nabla\mathfrak{m}_n(t)\|_2^2 + \alpha_1 \sup_{t\in[0,T]}\|\mathcal{A}\mathfrak{m}_n(t)\|_2^2\right] \\
&\leq C\mathbb{E}\left[\exp\left\{CT + C\int_0^T \|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^2 \mathrm{d}s\right\} \cdot \|\boldsymbol{\psi}(s)\|^2_{\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2)}\right] \\
&\leq C\left\{\mathbb{E}\left[\exp\left\{\frac{pC}{p-2}(T + \int_0^T \|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^2 \mathrm{d}s)\right\}\right]\right\}^{\frac{p-2}{p}} \cdot \|\boldsymbol{\psi}(s)\|^2_{\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2))} \\
&< +\infty,
\end{aligned} \tag{169}
$$

where we have used (33) and the fact that $\boldsymbol{\psi} \in \mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p > 2$). Therefore, by Banach-Alaoglu theorem, and uniqueness of weak and weak* limits, we have (along a subsequence)

$$
\mathfrak{m}_n \xrightarrow{w} \mathfrak{m} \quad \text{in} \quad \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\mathrm{D}(\mathcal{A}))). \tag{170}
$$

This completes the proof of Theorem 4.2. □

5. Gâteaux differentiability of the control-to-state mapping.

This section analyzes the differentiability of the control-to-state mapping. Specifically, we demonstrate that the solution to the linearized equation corresponds to the Gâteaux derivative of the control-to-state mapping. In order to prove such result, we need the following stability result:

Proposition 5.1. *Let $\boldsymbol{v}_1, \boldsymbol{v}_2$ be two solutions to system (31) in the sense of Definition 3.1 satisfying (33) associated with forcing $\boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))$ and same initial data $\boldsymbol{v}_0 \in \mathrm{D}(\mathcal{A}^{\frac{3}{2}})$, respectively. Then, there exists a constant K depending only on the given data ν, α_1, α_2 and β such that*

$$
\sup_{t\in[0,T]}\|\mathcal{A}\mathfrak{v}(t)\|_2^2 \leq Ce^{C\int_0^T[1+\|\boldsymbol{v}_1(t)\|_{\dot{\mathbb{H}}^3}^2+\|\boldsymbol{v}_2(t)\|_{\dot{\mathbb{H}}^3}^2]\mathrm{d}s}\|\boldsymbol{f}_1 - \boldsymbol{f}_2\|^2_{\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2)}, \quad \mathbb{P}\text{-}a.s. \tag{171}
$$

Proof. Defining $\mathfrak{v} := \boldsymbol{v}_1 - \boldsymbol{v}_2$, then the following equation holds

$$
\begin{cases}
\partial_t \Upsilon(\mathfrak{v}(t)) = -\left[\mathcal{G}(\boldsymbol{v}_1(t),\mathbf{0}) - \mathcal{G}(\boldsymbol{v}_2(t),\mathbf{0})\right] + \boldsymbol{f}_1 - \boldsymbol{f}_2, \\
\mathfrak{v}(0) = \boldsymbol{v}_{0,1} - \boldsymbol{v}_{0,2},
\end{cases} \tag{172}
$$

for a.e. $t \in [0,T]$ and $\mathbb{P}$-a.s. in $\mathbb{V}'$, where $\mathcal{G}(\cdot,\cdot)$ is given by (71). In view of $\Upsilon(\boldsymbol{v}_1 - \boldsymbol{z})$, $\Upsilon(\boldsymbol{v}_2 - \boldsymbol{z}) \in \mathrm{L}^2(0,T;\mathbb{V})$ (see (127)) and $\partial_t\Upsilon(\boldsymbol{v}_1 - \boldsymbol{z})$, $\partial_t\Upsilon(\boldsymbol{v}_1 - \boldsymbol{z}) \in \mathrm{L}^2(0,T;\mathbb{V}')$ (see (69)), we infer that $\Upsilon(\mathfrak{v}) \in \mathrm{L}^2(0,T;\mathbb{V})$ and $\partial_t\Upsilon(\mathfrak{v}) \in \mathrm{L}^2(0,T;\mathbb{V}')$. Therefore, an application of [31, Lemma 1.2, p.176] implies $\Upsilon(\mathfrak{v}) \in \mathrm{C}([0,T];\mathbb{H})$, the real-valued function $t \mapsto \|\Upsilon(\mathfrak{v}(t))\|_2^2$ is absolutely continuous and the following equality is satisfied:

$$
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\Upsilon(\mathfrak{v}(t))\|_2^2 = \langle\partial_t\Upsilon(\mathfrak{v}(t)), \Upsilon(\mathfrak{v}(t))\rangle, \qquad \text{for a.e. } t \in [0,T],
$$

or equivalently

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\|\mathfrak{v}(t)\|_2^2+2\alpha_1\|\nabla\mathfrak{v}(t)\|_2^2+\alpha_1^2\|\mathcal{A}\mathfrak{v}(t)\|_2^2\right]\\
&=-\left\langle\mathcal{G}(\boldsymbol{v}_1(t),\mathbf{0})-\mathcal{G}(\boldsymbol{v}_2(t),\mathbf{0}),\Upsilon(\mathfrak{v}(t))\right\rangle+\underbrace{\left\langle\boldsymbol{f}_1(t)-\boldsymbol{f}_2(t),\Upsilon(\mathfrak{v}(t))\right\rangle}_{=:V_0},
\end{aligned}\tag{173}$$

for a.e. $t\in[0,T]$, $\mathbb{P}$-a.s. Our next aim is to estimate the right hand side of (173) appropriately. Using integration by parts, (11), (13), we infer

$$\begin{aligned}
&-\left\langle\mathcal{G}(\boldsymbol{v}_1,\mathbf{0})-\mathcal{G}(\boldsymbol{v}_2,\mathbf{0}),\Upsilon(\mathfrak{v})\right\rangle\\
&=-\nu\|\nabla\mathfrak{v}\|_2^2-\nu\alpha_1\|\mathcal{A}\mathfrak{v}\|_2^2-\frac{\beta}{2}\int_{\mathbb{T}^2}|\mathrm{A}(\mathfrak{v}(x))|^2(|\mathrm{A}(\boldsymbol{v}_1(x))|^2+|\mathrm{A}(\boldsymbol{v}_2(x))|^2)\mathrm{d}x\\
&\quad+\underbrace{\beta\alpha_1\left\langle\mathrm{div}\{[\mathrm{A}(\mathfrak{v}):\mathrm{A}(\boldsymbol{v}_1)]\mathrm{A}(\boldsymbol{v}_1)+[\mathrm{A}(\mathfrak{v}):\mathrm{A}(\boldsymbol{v}_2)]\mathrm{A}(\boldsymbol{v}_1)+|\mathrm{A}(\boldsymbol{v}_2)|^2\mathrm{A}(\mathfrak{v})\},\mathcal{A}\mathfrak{v}\right\rangle}_{=:V_1}\\
&\quad-\underbrace{[b(\mathfrak{v},\Upsilon(\boldsymbol{v}_1),\Upsilon(\mathfrak{v}))-b(\Upsilon(\mathfrak{v}),\Upsilon(\boldsymbol{v}_1),\mathfrak{v})+b(\Upsilon(\mathfrak{v}),\boldsymbol{v}_2,\Upsilon(\mathfrak{v}))]}_{=:V_2}\\
&\quad-\underbrace{(\alpha_1+\alpha_2)\left\langle\mathrm{div}[\mathrm{A}(\boldsymbol{v}_1)\mathrm{A}(\mathfrak{v})+\mathrm{A}(\mathfrak{v})\mathrm{A}(\boldsymbol{v}_2)],\Upsilon(\mathfrak{v})\right\rangle}_{=:V_3}.
\end{aligned}\tag{174}$$

Using Hölder's, Sobolov's and Young's inequalities, we have

$$\begin{aligned}
|V_1|&=\left|\int_{\mathbb{T}^2}\mathfrak{D}^2\mathfrak{v}\cdot\mathfrak{D}^1\boldsymbol{v}_1\cdot\mathfrak{D}^1\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x+\int_{\mathbb{T}^2}\mathfrak{D}^1\mathfrak{v}\cdot\mathfrak{D}^2\boldsymbol{v}_1\cdot\mathfrak{D}^1\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\right.\\
&\quad+\int_{\mathbb{T}^2}\mathfrak{D}^2\mathfrak{v}\cdot\mathfrak{D}^1\boldsymbol{v}_2\cdot\mathfrak{D}^1\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x+\int_{\mathbb{T}^2}\mathfrak{D}^1\mathfrak{v}\cdot\mathfrak{D}^2\boldsymbol{v}_2\cdot\mathfrak{D}^1\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\\
&\quad+\int_{\mathbb{T}^2}\mathfrak{D}^1\mathfrak{v}\cdot\mathfrak{D}^1\boldsymbol{v}_2\cdot\mathfrak{D}^2\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x+\int_{\mathbb{T}^2}\mathfrak{D}^2\mathfrak{v}\cdot\mathfrak{D}^1\boldsymbol{v}_2\cdot\mathfrak{D}^1\boldsymbol{v}_2\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\\
&\quad\left.+\int_{\mathbb{T}^2}\mathfrak{D}^1\mathfrak{v}\cdot\mathfrak{D}^2\boldsymbol{v}_2\cdot\mathfrak{D}^1\boldsymbol{v}_2\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\right|\\
&\le C\|\mathfrak{D}^1\boldsymbol{v}_1\|_\infty^2\|\mathfrak{D}^2\mathfrak{v}\|_2^2+C\|\mathfrak{D}^1\mathfrak{v}\|_4\|\mathfrak{D}^2\boldsymbol{v}_1\|_4\|\mathfrak{D}^1\boldsymbol{v}_1\|_\infty\|\mathfrak{D}^2\mathfrak{v}\|_2\\
&\quad+C\|\mathfrak{D}^1\boldsymbol{v}_2\|_\infty\|\mathfrak{D}^1\boldsymbol{v}_1\|_\infty\|\mathfrak{D}^2\mathfrak{v}\|_2^2+C\|\mathfrak{D}^1\mathfrak{v}\|_4\|\mathfrak{D}^2\boldsymbol{v}_2\|_4\|\mathfrak{D}^1\boldsymbol{v}_1\|_\infty\|\mathfrak{D}^2\mathfrak{v}\|_2\\
&\quad+C\|\mathfrak{D}^1\mathfrak{v}\|_4\|\mathfrak{D}^1\boldsymbol{v}_2\|_\infty\|\mathfrak{D}^2\boldsymbol{v}_1\|_4\|\mathfrak{D}^2\mathfrak{v}\|_2+C\|\mathfrak{D}^1\boldsymbol{v}_2\|_\infty^2\|\mathfrak{D}^2\mathfrak{v}\|_2^2\\
&\quad+C\|\mathfrak{D}^1\mathfrak{v}\|_4\|\mathfrak{D}^2\boldsymbol{v}_2\|_4\|\mathfrak{D}^1\boldsymbol{v}_2\|_\infty\|\mathfrak{D}^2\mathfrak{v}\|_2\\
&\le C\big[\|\boldsymbol{v}_1\|_{\mathbb{H}^3}^2+\|\boldsymbol{v}_2\|_{\mathbb{H}^3}^2\big]\|\mathcal{A}\mathfrak{v}\|_2^2,
\end{aligned}\tag{175}$$

$$\begin{aligned}
|V_2+V_3|&=\left|\int_{\mathbb{T}^2}\mathfrak{D}^0\mathfrak{v}\cdot\mathfrak{D}^3\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x+\int_{\mathbb{T}^2}\mathfrak{D}^2\mathfrak{v}\cdot\mathfrak{D}^1\boldsymbol{v}_2\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\right.\\
&\quad+\int_{\mathbb{T}^2}\mathfrak{D}^2\boldsymbol{v}_1\cdot\mathfrak{D}^1\mathfrak{v}\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x+\int_{\mathbb{T}^2}\mathfrak{D}^1\boldsymbol{v}_1\cdot\mathfrak{D}^2\mathfrak{v}\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\\
&\quad\left.+\int_{\mathbb{T}^2}\mathfrak{D}^1\mathfrak{v}\cdot\mathfrak{D}^2\boldsymbol{v}_2\cdot\mathfrak{D}^2\mathfrak{v}\mathrm{d}x\right|\\
&\le C\|\mathfrak{v}\|_\infty\|\mathfrak{D}^3\boldsymbol{v}_1\|_2\|\mathfrak{D}^2\mathfrak{v}\|_2+C\|\mathfrak{D}^1\boldsymbol{v}_2\|_\infty\|\mathfrak{D}^2\mathfrak{v}\|_2^2\\
&\quad+C\|\mathfrak{D}^2\boldsymbol{v}_1\|_4\|\mathfrak{D}^1\mathfrak{v}\|_4\|\mathfrak{D}^2\mathfrak{v}\|_2+C\|\mathfrak{D}^1\boldsymbol{v}_1\|_\infty\|\mathfrak{D}^2\mathfrak{v}\|_2^2\\
&\quad+C\|\mathfrak{D}^1\mathfrak{v}\|_4\|\mathfrak{D}^2\boldsymbol{v}_1\|_4\|\mathfrak{D}^2\mathfrak{v}\|_2\\
&\le C\big[1+\|\boldsymbol{v}_1\|_{\mathbb{H}^3}^2+\|\boldsymbol{v}_2\|_{\mathbb{H}^3}^2\big]\|\mathcal{A}\mathfrak{v}\|_2^2,
\end{aligned}\tag{176}$$

$$|V_0| \leq C\|\boldsymbol{f}_1 - \boldsymbol{f}_2\|_2^2 + C\|\mathcal{A}\mathfrak{v}\|_2^2. \tag{177}$$

In view of (173)-(177), we write

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}\left[\|\mathfrak{v}(t)\|_2^2 + 2\alpha_1\|\nabla\mathfrak{v}(t)\|_2^2 + \alpha_1^2\|\mathcal{A}\mathfrak{v}(t)\|_2^2\right] \\
&\leq C\|\boldsymbol{f}_1(t) - \boldsymbol{f}_2(t)\|_2^2 + C\left[1 + \|\boldsymbol{v}_1(t)\|_{\dot{\mathbb{H}}^3}^2 + \|\boldsymbol{v}_2(t)\|_{\dot{\mathbb{H}}^3}^2\right] \\
&\quad \times \left[\|\mathfrak{v}(t)\|_2^2 + 2\alpha_1\|\nabla\mathfrak{v}(t)\|_2^2 + \alpha_1^2\|\mathcal{A}\mathfrak{v}(t)\|_2^2\right],
\end{aligned} \tag{178}$$

for a.e. $t \in [0,T]$ and $\mathbb{P}$-a.s. Hence, applying Gronwall's inequality, we complete the proof. □

Let us now prove the main result of this section in the following proposition.

Proposition 5.2. *Let* $\boldsymbol{\psi} \in \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))) \cap \mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ *(for some* $p>4$*),* $\boldsymbol{v}_0 \in \mathrm{D}(\mathcal{A}^{\frac{3}{2}})$ *and* $\boldsymbol{f} \in \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))$*. Let* $\boldsymbol{v}$ *and* $\boldsymbol{v}_\rho$ *be two solutions of system* (31) *in the sense of Definition 3.1 with same initial data* $\boldsymbol{v}_0$ *and corresponding to external forcing* $\boldsymbol{f}$ *and* $\boldsymbol{f}_\rho := \boldsymbol{f} + \rho\boldsymbol{\psi}$*,* $\rho \in (0,1)$*, respectively, then*

$$\chi_\rho := \frac{\boldsymbol{v}_\rho - \boldsymbol{v}}{\rho} - \mathfrak{m}, \quad \textit{satisfies} \quad \lim_{\rho\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}\left[\|\chi_\rho(t)\|_2^2 + \alpha_1\|\nabla\chi_\rho(t)\|_2^2\right]\right] = 0, \tag{179}$$

where $\mathfrak{m}$ *is the solution to system* (130) *in the sense of Definition of 4.1.*

Proof. Since $\boldsymbol{v}_\rho$ and $\boldsymbol{v}$ are the solution of system (31) corresponding to $\boldsymbol{f}+\rho\boldsymbol{\psi}$ and $\boldsymbol{f}$, respectively, and $\mathfrak{m}$ is the solution of system (130), therefore $\chi_\rho := \mathfrak{m}_\rho - \mathfrak{m} := \frac{\boldsymbol{v}_\rho - \boldsymbol{v}}{\rho} - \mathfrak{m}$ satisfies

$$\begin{cases}
\partial_t\Upsilon(\chi_\rho) + \nu\mathcal{A}\chi_\rho + \rho\mathcal{P}[(\mathfrak{m}_\rho\cdot\nabla)\Upsilon(\mathfrak{m}_\rho)] + \mathcal{P}[(\boldsymbol{v}\cdot\nabla)\Upsilon(\chi_\rho)] \\
+ \mathcal{P}[(\chi_\rho\cdot\nabla)\Upsilon(\boldsymbol{v})] + \rho\mathcal{P}\left[\sum_{j=1}^2[\Upsilon(\mathfrak{m}_\rho)]^j\nabla\mathfrak{m}_\rho^j\right] \\
+ \mathcal{P}\left[\sum_{j=1}^2\{[\Upsilon(\boldsymbol{v})]^j\nabla\chi_\rho^j + [\Upsilon(\chi_\rho)]^j\nabla\boldsymbol{v}^j\}\right] - (\alpha_1+\alpha_2)\rho\mathcal{P}\mathrm{div}[\mathrm{A}(\mathfrak{m}_\rho)\mathrm{A}(\mathfrak{m}_\rho)] \\
- (\alpha_1+\alpha_2)\mathcal{P}\mathrm{div}[\mathrm{A}(\chi_\rho)\mathrm{A}(\boldsymbol{v}) + \mathrm{A}(\boldsymbol{v})\mathrm{A}(\chi_\rho)] - \beta\rho^2\mathcal{P}\mathrm{div}[|\mathrm{A}(\mathfrak{m}_\rho)|^2\mathrm{A}(\mathfrak{m}_\rho)] \\
- \beta\rho\mathcal{P}\mathrm{div}[|\mathrm{A}(\mathfrak{m}_\rho)|^2\mathrm{A}(\boldsymbol{v})] - 2\beta\rho\mathcal{P}\mathrm{div}[(\mathrm{A}(\mathfrak{m}_\rho):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\mathfrak{m}_\rho)] \\
- 2\beta\mathcal{P}\mathrm{div}[(\mathrm{A}(\chi_\rho):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})] - \beta\mathcal{P}\mathrm{div}[|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\chi_\rho)] = \mathbf{0}, \\
\chi_\rho(0) = \mathbf{0},
\end{cases} \tag{180}$$

for a.e. $t\in[0,T]$ and $\mathbb{P}$-a.s., in $\mathbb{V}'$. Taking the inner product by χ_ρ to the equation $(180)_1$, and using the integration by parts, (11) and (13), we have

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}\left[\|\chi_\rho(t)\|_2^2 + \alpha_1\|\nabla\chi_\rho(t)\|_2^2\right] + \nu\|\nabla\chi_\rho(t)\|_2^2 + \frac{\beta}{2}\||\mathrm{A}(\boldsymbol{v}(t))|\mathrm{A}(\chi_\rho(t))\|_2^2 \\
&+ \beta\int_{\mathbb{T}^2}[\mathrm{A}(\boldsymbol{v}(x,t)):\mathrm{A}(\chi_\rho(x,t))]^2\mathrm{d}x \\
&= \underbrace{\rho b(\mathfrak{m}_\rho(t),\chi_\rho(t),\Upsilon(\mathfrak{m}_\rho(t)))}_{=:X_1} + \underbrace{\alpha_1 b(\boldsymbol{v}(t),\chi_\rho(t),\mathcal{A}\chi_\rho(t))}_{=:X_2}
\end{aligned}$$

$$
\begin{aligned}
&- \underbrace{\alpha_1 \rho b(\chi_\rho(t), \mathfrak{m}_\rho(t), \mathcal{A}\mathfrak{m}_\rho(t))}_{=:X_3} - \underbrace{b(\chi_\rho(t), \boldsymbol{v}(t), \Upsilon(\chi_\rho(t)))}_{=:X_4} \\
&\underbrace{- \frac{\alpha}{2} \int_{\mathbb{T}^2} [\mathrm{A}(\chi_\rho(x,t))\mathrm{A}(\boldsymbol{v}(x,t)) + \mathrm{A}(\boldsymbol{v}(x,t))\mathrm{A}(\chi_\rho(x,t))] : \mathrm{A}(\chi_\rho(x,t))\mathrm{d}x}_{=:X_5} \\
&\underbrace{- \frac{\alpha\rho}{2} \int_{\mathbb{T}^2} \mathrm{A}(\mathfrak{m}_\rho(x,t))\mathrm{A}(\mathfrak{m}_\rho(x,t)) : \mathrm{A}(\chi_\rho(x,t))\mathrm{d}x}_{=:X_6} \\
&\underbrace{- \frac{\beta\rho}{2} \int_{\mathbb{T}^2} |\mathrm{A}(\mathfrak{m}_\rho(x,t))|^2 [\mathrm{A}(\boldsymbol{v}_\rho(x,t)) : \mathrm{A}(\chi_\rho(x,t))]\mathrm{d}x}_{=:X_7} \\
&\underbrace{- \beta\rho \int_{\mathbb{T}^2} [\mathrm{A}(\mathfrak{m}_\rho(x,t)) : \mathrm{A}(\boldsymbol{v}(x,t))][\mathrm{A}(\mathfrak{m}_\rho(x,t)) : \mathrm{A}(\chi_\rho(x,t))]\mathrm{d}x}_{=:X_8}, \qquad (181)
\end{aligned}
$$

for a.e. $t \in [0,T]$ and $\mathbb{P}$-a.s. Let us now estimate each term of right hand side of (181) using Hölder's, Sobolev's and Young's inequalities as follows:

$$
\begin{aligned}
|X_1| &\leq C\rho \|\mathfrak{m}_\rho\|_\infty \|\nabla \chi_\rho\|_2 \|\mathcal{A}\mathfrak{m}_\rho\|_2 \\
&\leq C\rho \|\nabla \chi_\rho\|_2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^2 \\
&\leq C\|\nabla \chi_\rho\|_2^2 + C\rho^2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^4, \qquad (182) \\
|X_2| &= \left| \int_{\mathbb{T}^2} \mathfrak{D}^1 \boldsymbol{v} \cdot \mathfrak{D}^1 \chi_\rho \cdot \mathfrak{D}^1 \chi_\rho \mathrm{d}x \right| \\
&\leq C\|\nabla \boldsymbol{v}\|_\infty \|\nabla \chi_\rho\|_2^2 \\
&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\nabla \chi_\rho\|_2^2, \qquad (183) \\
|X_3| &\leq C\rho \|\chi_\rho\|_4 \|\nabla \mathfrak{m}_\rho\|_4 \|\mathcal{A}\mathfrak{m}_\rho\|_2 \\
&\leq C\rho \|\nabla \chi_\rho\|_2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^2 \\
&\leq C\|\nabla \chi_\rho\|_2^2 + C\rho^2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^4, \qquad (184) \\
|X_4| &= \left| \int_{\mathbb{T}^2} \mathfrak{D}^1 \chi_\rho \cdot \mathfrak{D}^1 \boldsymbol{v} \cdot \mathfrak{D}^1 \chi_\rho \mathrm{d}x + \int_{\mathbb{T}^2} \mathfrak{D}^0 \chi_\rho \cdot \mathfrak{D}^2 \boldsymbol{v} \cdot \mathfrak{D} \chi_\rho \mathrm{d}x \right| \\
&\leq C\|\nabla \chi_\rho\|_2 \|\nabla \boldsymbol{v}\|_\infty \|\nabla \chi_\rho\|_2 + C\|\chi_\rho\|_4 \|\mathfrak{D}^2 \boldsymbol{v}\|_4 \|\nabla \chi_\rho\|_2 \\
&\leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\nabla \chi_\rho\|_2^2, \qquad (185) \\
|X_5| &\leq C\|\nabla \boldsymbol{v}\|_\infty \|\nabla \chi_\rho\|_2^2 \leq C\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} \|\nabla \chi_\rho\|_2^2, \qquad (186) \\
|X_6| &\leq C\rho \|\nabla \mathfrak{m}_\rho\|_4^2 \|\nabla \chi_\rho\|_2 \\
&\leq C\rho \|\nabla \chi_\rho\|_2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^2 \\
&\leq C\|\nabla \chi_\rho\|_2^2 + C\rho^2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^4, \qquad (187) \\
|X_7 + X_8| &\leq C\rho [\|\nabla \boldsymbol{v}\|_\infty + \|\nabla \boldsymbol{v}_\rho\|_\infty] \|\nabla \mathfrak{m}_\rho\|_4^2 \|\nabla \chi_\rho\|_2 \\
&\leq C\rho [\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3} + \|\boldsymbol{v}_\rho\|_{\dot{\mathbb{H}}^3}] \|\nabla \mathfrak{m}_\rho\|_4^2 \|\nabla \chi_\rho\|_2 \\
&\leq C[\|\boldsymbol{v}\|_{\dot{\mathbb{H}}^3}^2 + \|\boldsymbol{v}_\rho\|_{\dot{\mathbb{H}}^3}^2] \|\nabla \chi_\rho\|_2^2 + C\rho^2 \|\mathcal{A}\mathfrak{m}_\rho\|_2^4. \qquad (188)
\end{aligned}
$$

Combining (181)-(188), we reach at

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\bigg[\|\chi_\rho(t)\|_2^2+\alpha_1\|\nabla\chi_\rho(t)\|_2^2\bigg]\\
&\leq C\big\{1+\|\boldsymbol{v}(t)\|_{\dot{\mathbb{H}}^3}^2+\|\boldsymbol{v}_\rho(t)\|_{\dot{\mathbb{H}}^3}^2\big\}\bigg[\|\chi_\rho(t)\|_2^2+\alpha_1\|\nabla\chi_\rho(t)\|_2^2\bigg]\\
&\quad+C\rho^2\|\mathcal{A}\mathfrak{m}_\rho(t)\|_2^4. \qquad (189)
\end{aligned}$$

for a.e. $t\in[0,T]$ and $\mathbb{P}$-a.s. Applying Gronwall's inequality to (189) and using (171), we obtain

$$\begin{aligned}
\|\chi_\rho(t)\|_2^2+\alpha_1\|\nabla\chi_\rho(t)\|_2^2&\leq C\rho^2e^{C\int_0^t[1+\|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^2+\|\boldsymbol{v}_\rho(s)\|_{\dot{\mathbb{H}}^3}^2]\mathrm{d}s}\cdot\|\mathcal{A}\mathfrak{m}_\rho(t)\|_2^4\\
&\leq C\rho^2e^{C\int_0^t[1+\|\boldsymbol{v}(s)\|_{\dot{\mathbb{H}}^3}^2+\|\boldsymbol{v}_\rho(s)\|_{\dot{\mathbb{H}}^3}^2]\mathrm{d}s}\cdot\|\boldsymbol{\psi}\|_{\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2)}^4. \qquad (190)
\end{aligned}$$

for all $t\in[0,T]$ and $\mathbb{P}$-a.s. Taking expectation of (190), and using Hölder's inequality, (33) and the fact that $\boldsymbol{\psi}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>4$), we complete the proof. ☐

5.1. **Variation of the cost functional (26).** As a consequence of Proposition 5.2, we get the following result on the variation for the cost functional (26).

Proposition 5.3. *Let $\boldsymbol{\psi}\in\mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))\cap\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>4$), $\boldsymbol{v}_0\in\mathrm{D}(\mathcal{A}^{\frac{3}{2}})$ and $\boldsymbol{f}\in\mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2)))$. Let $\boldsymbol{v},\boldsymbol{v}_\rho$ be the solutions of system (31) corresponding to external forcing $\boldsymbol{f}$ and $\boldsymbol{f}_\rho:=\boldsymbol{f}+\rho\boldsymbol{\psi}$, $\rho\in(0,1)$, respectively, in the sense of Definition 3.1 and $\mathfrak{m}$ be the solution of system (130) in the sense of Definition 4.1. Then, we have*

$$\begin{aligned}
&\mathrm{J}(\boldsymbol{f}_\rho,\boldsymbol{v}_\rho)-\mathrm{J}(\boldsymbol{f},\boldsymbol{v})\\
&=\rho\,\mathbb{E}\left[\int_0^T\{(\nabla_{\boldsymbol{f}}\mathfrak{L}(t,\boldsymbol{f}(t),\boldsymbol{v}(t)),\boldsymbol{\psi}(t))+(\nabla_{\boldsymbol{v}}\mathfrak{L}(t,\boldsymbol{f}(t),\boldsymbol{v}(t)),\mathfrak{m}(t))\}\mathrm{d}t\right]+o(\rho).
\end{aligned}$$

6. **Adjoint system.** The main goal of this section is to show that, for the unique solution $\boldsymbol{v}$ to the state equation (31), as defined in Definition 3.4, satisfying (33) and $\boldsymbol{g}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>2$), the following adjoint system is well-posed:

$$\begin{cases}
-\partial_t(\mathfrak{p}-\alpha_1\Delta\mathfrak{p})=\boldsymbol{g}-\nabla\pi^*+\nu\Delta\mathfrak{p}+(\mathrm{I}-\alpha_1\Delta)[(\boldsymbol{v}\cdot\nabla)\mathfrak{p}]-\displaystyle\sum_{j=1}^2\mathfrak{p}^j\nabla(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v})^j\\
\qquad-(\mathrm{I}-\alpha_1\Delta)[(\mathfrak{p}\cdot\nabla)\boldsymbol{v}]+(\mathfrak{p}\cdot\nabla)(\boldsymbol{v}-\alpha_1\Delta\boldsymbol{v})\\
\qquad+(\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{p})+\mathrm{A}(\mathfrak{p})\mathrm{A}(\boldsymbol{v}))+\beta[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{p}))]\\
\qquad+2\beta\mathrm{div}[(\mathrm{A}(\mathfrak{p}):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})], & \text{in } \mathbb{T}^2\times(0,T],\\
\mathrm{div}\,\mathfrak{p}=0, & \text{in } \mathbb{T}^2\times[0,T],\\
\mathfrak{p}(x,T)=\mathbf{0}, & \text{in } \mathbb{T}^2,
\end{cases} \qquad (191)$$

with periodic boundary conditions, where $\mathfrak{p}=(\mathfrak{p}^1,\mathfrak{p}^2)$ and π^* are unknown vector and scalar fields, receptively, and $\mathrm{A}(\mathfrak{p})=\nabla\mathfrak{p}+(\nabla\mathfrak{p})^T$. Taking the projection $\mathcal{P}$ to the system (191), we obtain

$$\begin{cases} -\partial_t \Upsilon(\mathfrak{p}) = \mathcal{P}\boldsymbol{g} - \nu\mathcal{A}\mathfrak{p} + \mathcal{P}[(\mathrm{I}-\alpha_1\Delta)[(\boldsymbol{v}\cdot\nabla)\mathfrak{p}]] - \mathcal{P}\left[\sum_{j=1}^{2}\mathfrak{p}^j\nabla[\Upsilon(\boldsymbol{v})]^j\right] \\ \qquad\qquad - \mathcal{P}\left[(\mathrm{I}-\alpha_1\Delta)[(\mathfrak{p}\cdot\nabla)\boldsymbol{v}]\right] + \mathcal{P}[(\mathfrak{p}\cdot\nabla)\Upsilon(\boldsymbol{v})] \\ \qquad\qquad + (\alpha_1+\alpha_2)\mathcal{P}[\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{p}) + \mathrm{A}(\mathfrak{p})\mathrm{A}(\boldsymbol{v}))] + \beta\mathcal{P}[\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{p}))] \\ \qquad\qquad + 2\beta\mathcal{P}[\mathrm{div}[(\mathrm{A}(\mathfrak{p}):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})]], \\ \mathfrak{p}(T) = \boldsymbol{0}, \end{cases} \tag{192}$$

in $\mathbb{V}'$. We obtain the well-posedness of system (192) in the following sense:

Definition 6.1. A stochastic process $\{\mathfrak{p}(t)\}_{t\in[0,T]}$ with trajectories in $\mathrm{C}([0,T];\mathbb{V})\cap\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$, $\partial_t\mathfrak{p}\in\mathrm{L}^2(0,T;\mathbb{V})$ and $\mathfrak{p}(T)=\boldsymbol{0}$, is called a *solution* to system (192), if for $\boldsymbol{g}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>2$), it satisfies for any $\phi\in\mathbb{V}$,

$$\begin{aligned} &-\langle\Upsilon(\mathfrak{p}(t)),\phi\rangle \\ &= \Big\langle \boldsymbol{g}(t) - \nu\mathcal{A}\mathfrak{p}(t) + (\mathrm{I}-\alpha_1\Delta)[(\boldsymbol{v}\cdot\nabla)\mathfrak{p}] - \sum_{j=1}^{2}\mathfrak{p}^j\nabla[\Upsilon(\boldsymbol{v})]^j - (\mathrm{I}-\alpha_1\Delta)[(\mathfrak{p}\cdot\nabla)\boldsymbol{v}] \\ &\quad + (\mathfrak{p}\cdot\nabla)\Upsilon(\boldsymbol{v}) + (\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v}(t))\mathrm{A}(\mathfrak{p}(t)) + \mathrm{A}(\mathfrak{p}(t))\mathrm{A}(\boldsymbol{v}(t))) \\ &\quad + \beta\mathrm{div}(|\mathrm{A}(\boldsymbol{v}(t))|^2\mathrm{A}(\mathfrak{p}(t))) + 2\beta\mathrm{div}((\mathrm{A}(\mathfrak{p}(t)):\mathrm{A}(\boldsymbol{v}(t)))\mathrm{A}(\boldsymbol{v}(t))),\phi\Big\rangle, \end{aligned} \tag{193}$$

for a.e. $t\in[0,T]$ and $\mathbb{P}$-a.s.

Next theorem is the main result of this section which provides the existence and uniqueness of global solutions to system (192) in the sense of Definition 6.1.

Theorem 6.2. *There exists a unique solution* $\mathfrak{p}$ *to system* (192) *in the sense of Definition 6.1.*

Proof of Theorem 6.2. Notice that $\mathfrak{p}$ is the solution of (191) if and only if $\mathfrak{q}(t)=\mathfrak{p}(T-t)$ is the solution of the following initial value problem with $\bar{\boldsymbol{v}}(t)=\boldsymbol{v}(T-t),\bar{\boldsymbol{g}}(t)=\boldsymbol{g}(T-t)$ and $\bar{\pi}^*(t)=\pi^*(T-t)$

$$\begin{cases} \partial_t(\mathfrak{q}-\alpha_1\Delta\mathfrak{q}) = \bar{\boldsymbol{g}} - \nabla\bar{\pi}^* + \nu\Delta\mathfrak{q} + (\mathrm{I}-\alpha_1\Delta)[(\bar{\boldsymbol{v}}\cdot\nabla)\mathfrak{q}] - \sum_{j=1}^{2}\mathfrak{q}^j\nabla(\bar{\boldsymbol{v}}-\alpha_1\Delta\bar{\boldsymbol{v}})^j \\ \qquad\qquad - (\mathrm{I}-\alpha_1\Delta)[(\mathfrak{q}\cdot\nabla)\bar{\boldsymbol{v}}] + (\mathfrak{q}\cdot\nabla)(\bar{\boldsymbol{v}}-\alpha_1\Delta\bar{\boldsymbol{v}}) \\ \qquad\qquad + (\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\bar{\boldsymbol{v}})\mathrm{A}(\mathfrak{q}) + \mathrm{A}(\mathfrak{q})\mathrm{A}(\bar{\boldsymbol{v}})) + \beta[\mathrm{div}(|\mathrm{A}(\bar{\boldsymbol{v}})|^2\mathrm{A}(\mathfrak{q}))] \\ \qquad\qquad + 2\beta\mathrm{div}[(\mathrm{A}(\mathfrak{q}):\mathrm{A}(\bar{\boldsymbol{v}}))\mathrm{A}(\bar{\boldsymbol{v}})], & \text{in } \mathbb{T}^2\times(0,T], \\ \mathrm{div}\,\mathfrak{q} = 0, & \text{in } \mathbb{T}^2\times[0,T], \\ \mathfrak{q}(x,0) = \boldsymbol{0}, & \text{in } \mathbb{T}^2, \end{cases} \tag{194}$$

with periodic boundary conditions. According to the Definition 6.1, $\mathfrak{q}$ is the solution of system (194) if $\mathfrak{q}\in\mathrm{C}([0,T];\mathbb{V})\cap\mathrm{L}^\infty(0,T;\mathrm{D}(\mathcal{A}))$ with $\partial_t\mathfrak{q}\in\mathrm{L}^2(0,T;\mathbb{V}')$, $\mathfrak{q}(0)=\boldsymbol{0}$ and, in addition, the following equality holds, for all $\boldsymbol{\phi}\in\mathbb{V}$,

$$\begin{aligned} &\langle\partial_t\Upsilon(\mathfrak{q}),\boldsymbol{\phi}\rangle \\ &= \Big\langle\bar{\boldsymbol{g}} - \nu\mathcal{A}\mathfrak{q} + (\mathrm{I}-\alpha_1\Delta)[(\bar{\boldsymbol{v}}\cdot\nabla)\mathfrak{q}] - \sum_{j=1}^{2}\mathfrak{q}^j\nabla[\Upsilon(\bar{\boldsymbol{v}})]^j - (\mathrm{I}-\alpha_1\Delta)[(\mathfrak{q}\cdot\nabla)\bar{\boldsymbol{v}}] \end{aligned}$$

$$+ (\mathfrak{q}\cdot\nabla)\Upsilon(\bar{\boldsymbol{v}}) + (\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\bar{\boldsymbol{v}})\mathrm{A}(\mathfrak{q}) + \mathrm{A}(\mathfrak{q})\mathrm{A}(\bar{\boldsymbol{v}})) + \beta\mathrm{div}(|\mathrm{A}(\bar{\boldsymbol{v}})|^2\mathrm{A}(\mathfrak{q}))$$
$$+ 2\beta\mathrm{div}((\mathrm{A}(\mathfrak{q}) : \mathrm{A}(\bar{\boldsymbol{v}}))\mathrm{A}(\bar{\boldsymbol{v}})), \boldsymbol{\phi}\rangle, \tag{195}$$

for a.e. $t\in[0,T]$ and $\mathbb{P}$-a.s.

Let us consider the following approximate equation for system (195) on the finite-dimensional space $\mathbb{H}_n$:

$$\begin{cases}\partial_t\Upsilon(\mathfrak{q}_n(t)) = \mathrm{P}_n\mathcal{P}\bar{\boldsymbol{g}}(t) - \nu\mathrm{P}_n\mathcal{A}\mathfrak{q}_n(t) - \mathrm{P}_n\mathcal{P}[(\mathrm{I}-\alpha_1\Delta)[(\bar{\boldsymbol{v}}(t)\cdot\nabla)\mathfrak{q}_n(t)]] \\ \qquad - \mathrm{P}_n\mathcal{P}\left[\displaystyle\sum_{j=1}^{2}(\mathfrak{q}_n)^j\nabla[\Upsilon(\bar{\boldsymbol{v}}(t))]^j\right] - \mathrm{P}_n\mathcal{P}[(\mathrm{I}-\alpha_1\Delta)[(\mathfrak{q}_n(t)\cdot\nabla)\bar{\boldsymbol{v}}(t)]] \\ \qquad + \mathrm{P}_n\mathcal{P}[(\mathfrak{q}_n(t)\cdot\nabla)\Upsilon(\bar{\boldsymbol{v}}(t))] \\ \qquad + (\alpha_1+\alpha_2)\mathrm{P}_n\mathcal{P}[\mathrm{div}(\mathrm{A}(\bar{\boldsymbol{v}}(t))\mathrm{A}(\mathfrak{q}_n(t)) + \mathrm{A}(\mathfrak{q}_n(t))\mathrm{A}(\bar{\boldsymbol{v}}(t)))] \\ \qquad + \beta\mathrm{P}_n\mathcal{P}[\mathrm{div}(|\mathrm{A}(\bar{\boldsymbol{v}}(t))|^2\mathrm{A}(\mathfrak{q}_n(t)))] \\ \qquad + 2\beta\mathrm{P}_n\mathcal{P}[\mathrm{div}[(\mathrm{A}(\mathfrak{q}_n(t)) : \mathrm{A}(\bar{\boldsymbol{v}}(t)))\mathrm{A}(\bar{\boldsymbol{v}}(t))]], \\ \mathfrak{q}_n(0) = \mathbf{0}. \end{cases} \tag{196}$$

Since system (194) closely resembles system (129), applying the same reasoning as in the proof of Theorem 4.2 establishes the existence and uniqueness of its solution. Consequently, this also guarantees the existence and uniqueness of a solution for system (191) with the stated regularity. Ultimately, one also obtain

$$\mathfrak{p}_n \xrightarrow{w} \mathfrak{p} \quad \text{in} \qquad \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\mathrm{D}(\mathcal{A}))). \tag{197}$$

This completes the proof. □

7. **Existence of optimal control and optimality condition.**

This section establishes a duality relationship between the solutions of the adjoint equation and the linearized equation, demonstrating the existence of a solution to the control problem. Using this duality, we further show that the control problem's solution satisfies the first-order necessary optimality criterion.

7.1. **Duality property.**

In the following proposition, we establish a duality relation between the solutions of the adjoint equation and the linearized equation.

Proposition 7.1. *Let $\boldsymbol{v}$ satisfies (33) and $\boldsymbol{g},\boldsymbol{\psi}\in\mathrm{L}^p(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p>2$). Then we have*

$$\mathbb{E}\left[\int_0^T(\boldsymbol{\psi}(t),\mathfrak{p}(t))\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T(\boldsymbol{g}(t),\mathfrak{m}(t))\mathrm{d}t\right],$$

where $\mathfrak{p}$ is the solution of (191) and $\mathfrak{m}$ is the solution of (129).

Proof. Let us consider the following approximate equations corresponding to systems (130) and (192) on the finite-dimensional space $\mathbb{H}_n$, respectively:

$$\begin{cases}(\partial_t\Upsilon(\mathfrak{m}_n),\boldsymbol{w}_n)=(\boldsymbol{\psi}-\nu\mathcal{A}\mathfrak{m}_n-(\boldsymbol{v}\cdot\nabla)\Upsilon(\mathfrak{m}_n)-(\mathfrak{m}_n\cdot\nabla)\Upsilon(\boldsymbol{v})-\displaystyle\sum_{j=1}^{2}[\Upsilon(\mathfrak{m}_n)]^j\nabla\boldsymbol{v}^j\\ \qquad-\displaystyle\sum_{j=1}^{2}[\Upsilon(\boldsymbol{v})]^j\nabla[\mathfrak{m}_n]^j+(\alpha_1+\alpha_2)[\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{m}_n)+\mathrm{A}(\mathfrak{m}_n)\mathrm{A}(\boldsymbol{v}))]\\ \qquad+\beta\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{m}_n))+2\beta\mathrm{div}((\mathrm{A}(\mathfrak{m}_n):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})),\boldsymbol{w}_n),\\ \mathfrak{m}_n(0)=\mathbf{0},\end{cases}\tag{198}$$

for a.e. $t\in[0,T]$ and for all $\boldsymbol{w}_n\in\mathbb{H}_n$, $\mathbb{P}$-a.s., and

$$\begin{cases}-(\partial_t\Upsilon(\mathfrak{p}_n),\boldsymbol{w}_n)=(\boldsymbol{g}-\nu\mathcal{A}\mathfrak{p}_n-(\mathrm{I}-\alpha_1\Delta)[(\boldsymbol{v}\cdot\nabla)\mathfrak{p}_n]-\displaystyle\sum_{j=1}^{2}[\mathfrak{p}_n]^j\nabla[\Upsilon(\boldsymbol{v})]^j\\ \qquad-(\mathrm{I}-\alpha_1\Delta)[(\mathfrak{p}_n\cdot\nabla)\boldsymbol{v}]+(\mathfrak{p}_n\cdot\nabla)\Upsilon(\boldsymbol{v})\\ \qquad+(\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{p}_n)+\mathrm{A}(\mathfrak{p}_n)\mathrm{A}(\boldsymbol{v}))\\ \qquad+\beta\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{p}_n))+2\beta\mathrm{div}((\mathrm{A}(\mathfrak{p}_n):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})),\boldsymbol{w}_n),\\ \mathfrak{p}_n(T)=\mathbf{0},\end{cases}\tag{199}$$

for a.e. $t\in[0,T]$ and for all $\boldsymbol{w}_n\in\mathbb{H}_n$, $\mathbb{P}$-a.s.

Setting $\boldsymbol{w}_n=\mathfrak{p}_n$ in $(198)_1$, we get

$$\begin{aligned}(\partial_t\Upsilon(\mathfrak{m}_n),\mathfrak{p}_n)&=(\boldsymbol{\psi}-\nu\mathcal{A}\mathfrak{m}_n-(\boldsymbol{v}\cdot\nabla)\Upsilon(\mathfrak{m}_n)-(\mathfrak{m}_n\cdot\nabla)\Upsilon(\boldsymbol{v})-\sum_{j=1}^{2}[\Upsilon(\mathfrak{m}_n)]^j\nabla\boldsymbol{v}^j\\&\quad-\sum_{j=1}^{2}[\Upsilon(\boldsymbol{v})]^j\nabla[\mathfrak{m}_n]^j+(\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{m}_n)+\mathrm{A}(\mathfrak{m}_n)\mathrm{A}(\boldsymbol{v}))\\&\quad+\beta\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{m}_n))+2\beta\mathrm{div}((\mathrm{A}(\mathfrak{m}_n):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})),\mathfrak{p}_n),\end{aligned}\tag{200}$$

for a.e. $t\in[0,T]$ and $\mathbb{P}$-a.s. Also, setting $\boldsymbol{w}_n=\mathfrak{m}_n$ in $(199)_1$, we get

$$\begin{aligned}&-(\partial_t\Upsilon(\mathfrak{p}_n),\mathfrak{m}_n)\\&=(\boldsymbol{g}-\nu\mathcal{A}\mathfrak{p}_n-(\mathrm{I}-\alpha_1\Delta)[(\boldsymbol{v}\cdot\nabla)\mathfrak{p}_n]-\sum_{j=1}^{2}[\mathfrak{p}_n]^j\nabla[\Upsilon(\boldsymbol{v})]^j-(\mathrm{I}-\alpha_1\Delta)[(\mathfrak{p}_n\cdot\nabla)\boldsymbol{v}]\\&\quad+(\mathfrak{p}_n\cdot\nabla)\Upsilon(\boldsymbol{v})+(\alpha_1+\alpha_2)\mathrm{div}(\mathrm{A}(\boldsymbol{v})\mathrm{A}(\mathfrak{p}_n)+\mathrm{A}(\mathfrak{p}_n)\mathrm{A}(\boldsymbol{v}))+\beta\mathrm{div}(|\mathrm{A}(\boldsymbol{v})|^2\mathrm{A}(\mathfrak{p}_n))\\&\quad+2\beta\mathrm{div}((\mathrm{A}(\mathfrak{p}_n):\mathrm{A}(\boldsymbol{v}))\mathrm{A}(\boldsymbol{v})),\mathfrak{m}_n),\end{aligned}\tag{201}$$

for a.e. $t\in[0,T]$ and $\mathbb{P}$-a.s. The differentiation rules give

$$\begin{aligned}&-(\partial_t\Upsilon(\mathfrak{p}_n(t)),\mathfrak{m}_n(t))\\&=(\partial_t\Upsilon(\mathfrak{m}_n(t)),\mathfrak{p}_n(t))-\partial_t\big[(\mathfrak{p}_n(t),\mathfrak{m}_n(t))+\alpha_1(\nabla\mathfrak{p}_n(t),\nabla\mathfrak{m}_n(t))\big].\end{aligned}\tag{202}$$

Integrating (202) with respect to t over $[0,T]$, and using that $\mathfrak{m}_n(0)=\mathfrak{p}_n(T)=0$, we obtain

$$-\int_0^T(\partial_t\Upsilon(\mathfrak{p}_n(t)),\mathfrak{m}_n(t))\mathrm{d}t=\int_0^T(\partial_t\Upsilon(\mathfrak{m}_n(t)),\mathfrak{p}_n(t))\mathrm{d}t,\tag{203}$$

$\mathbb{P}$-a.s. Substituting (201) and (200) in (203), integrating by parts and using the fact that $(A, B) = (A^T, B^T)$, for any $A, B \in \mathcal{M}_{2\times 2}(\mathbb{R})$, we obtain

$$\mathbb{E}\left[\int_0^T (\boldsymbol{\psi}(t), \mathfrak{p}_n(t))\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T (\boldsymbol{g}(t), \mathfrak{m}_n(t))\mathrm{d}t\right].$$

Therefore, in view of (170) and (197), taking the limit as $n \to \infty$, we complete the proof. ☐

Considering $\boldsymbol{g} = \nabla_{\boldsymbol{v}}\mathfrak{L}(\cdot, \boldsymbol{f}, \boldsymbol{v}) \in \mathrm{L}^p(\Omega; \mathrm{L}^2(0, T; \dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p > 2$) in Proposition 7.1, we obtain the following result.

Corollary 7.2. *Under the assumptions of Proposition 7.1, the following duality relation holds*

$$\mathbb{E}\left[\int_0^T (\boldsymbol{\psi}(t), \mathfrak{p}(t))\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{v}}\mathfrak{L}(t, \boldsymbol{f}, \boldsymbol{v}), \mathfrak{m}(t))\mathrm{d}t\right].$$

7.2. **Existence of an optimal control for (27).** Suppose that $\{\boldsymbol{f}_n, \boldsymbol{v}_n\}_{n\in\mathbb{N}}$ is a minimizing sequence, remark that $\{\boldsymbol{f}_n\}_{n\in\mathbb{N}}$ belongs to $\mathcal{X} = \mathrm{L}^2(\Omega; \mathrm{L}^2(0, T; \dot{\mathbb{H}}^1(\mathbb{T}^2)))$ $\cap \mathrm{L}^p(\Omega; \mathrm{L}^2(0, T; \dot{\mathbb{L}}^2(\mathbb{T}^2)))$ (for some $p > 4$) and has uniformly random bounded.

Since

$$\mathcal{F}_{ad} := \{\boldsymbol{f} \in \mathcal{X} : \|\boldsymbol{f}(\omega)\|_{\mathrm{L}^2(0,T;\dot{\mathbb{H}}^1(\mathbb{T}^2))} \leq K\varrho(\omega), \text{ for a.s. } \omega \in \Omega\},$$

there exists $\widetilde{\boldsymbol{f}} \in \mathcal{F}_{ad}$ such that

$$\boldsymbol{f}_n \rightharpoonup \widetilde{\boldsymbol{f}} \quad \text{ in } \mathrm{L}^2(0, T; \dot{\mathbb{H}}^1(\mathbb{T}^2)), \ \ \mathbb{P}\text{-a.s.} \tag{204}$$

and

$$\boldsymbol{f}_n \rightharpoonup \widetilde{\boldsymbol{f}} \quad \text{ in } \mathrm{L}^2(\Omega; \mathrm{L}^2(0, T; \dot{\mathbb{L}}^2(\mathbb{T}^2))). \tag{205}$$

According to our construction of unique solution $\boldsymbol{v}_n$ to system (31) (where $\boldsymbol{f}$ is replaced by $\boldsymbol{f}_n$), there exists a unique solution $\boldsymbol{u}_n$ to system (35) (where $\boldsymbol{f}$ is replaced by $\boldsymbol{f}_n$) such that (see Theorems 3.2 and 3.5)

$$\begin{gathered}\{\boldsymbol{u}_n\}_{n\in\mathbb{N}} \text{ is uniformly bounded in} \\ \mathrm{C}([0, T]; \mathrm{D}(\mathcal{A})) \cap \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}})) \cap \mathrm{H}^1(0, T; \mathrm{D}(\mathcal{A})), \ \mathbb{P}\text{-a.s.},\end{gathered}$$

and

$$\{\boldsymbol{u}_n\}_{n\in\mathbb{N}} \text{ is uniformly bounded in } \mathrm{L}^q(\Omega; \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}}))), \quad q \geq 2.$$

By Banach-Alaoglu theorem, there exists

$$\begin{gathered}\widetilde{\boldsymbol{u}} \in \mathrm{L}^q(\Omega; \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}}))), \ q \geq 2, \text{ and} \\ \widetilde{\boldsymbol{u}} \in \mathrm{C}([0, T]; \mathrm{D}(\mathcal{A})) \cap \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}})) \cap \mathrm{H}^1(0, T; \mathrm{D}(\mathcal{A})), \ \mathbb{P}\text{-a.s.}\end{gathered}$$

such that up to a subsequence (denoting by the same)

$$\begin{cases} \boldsymbol{u}_n \rightharpoonup \widetilde{\boldsymbol{u}} & \text{in } \mathrm{L}^2(\Omega; \mathrm{L}^2(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}}))), \\ \boldsymbol{u}_n \xrightharpoonup{*} \widetilde{\boldsymbol{u}} & \text{in } \mathrm{L}^\infty(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}})), \ \ \mathbb{P}\text{-a.s.}, \\ \boldsymbol{u}_n \rightharpoonup \widetilde{\boldsymbol{u}} & \text{in } \mathrm{L}^2(0, T; \mathrm{D}(\mathcal{A}^{\frac{3}{2}})), \ \ \mathbb{P}\text{-a.s.}, \\ \partial_t\boldsymbol{u}_n \rightharpoonup \partial_t\widetilde{\boldsymbol{u}} & \text{in } \mathrm{L}^2(0, T; \mathrm{D}(\mathcal{A})), \ \ \mathbb{P}\text{-a.s.} \end{cases} \tag{206}$$

By (206) and [32, Lemma 3.1, p. 404], we observe that $\widetilde{\boldsymbol{u}} \in \mathrm{C}([0,T],\mathrm{D}(\mathcal{A}))$ and therefore $\boldsymbol{u}_n(0)$ converges to $\widetilde{\boldsymbol{u}}(0)$ in $\mathrm{D}(\mathcal{A})$, which gives $\widetilde{\boldsymbol{u}}(0)=\boldsymbol{v}_0$. Now, a similar arguments as in the proof of Theorem 3.5 provide that $(\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}}:=\widetilde{\boldsymbol{u}}+\boldsymbol{z})$ solves (31).

Recall that $\mathrm{J}:\mathcal{F}_{ad}\to\mathbb{R}^+$ given by (26) is convex and continuous. From Proposition 2.2, (205) and (206), we have

$$\boldsymbol{f}_n \rightharpoonup \widetilde{\boldsymbol{f}} \quad \text{in } \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\dot{\mathbb{L}}^2(\mathbb{T}^2))) \text{ and } \boldsymbol{v}_n \rightharpoonup \widetilde{\boldsymbol{v}} \quad \text{in } \mathrm{L}^2(\Omega;\mathrm{L}^2(0,T;\mathbb{H})).$$

The weak lower semicontinuity of J ensures

$$\mathrm{J}(\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}}) \leq \liminf_{n} \mathrm{J}(\boldsymbol{f}_n,\boldsymbol{v}_n),$$

which gives that $(\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}})$ is an optimal pair.

7.3. **First-order necessary optimality condition for (27).** To derive the first-order necessary optimality conditions, we consider the reduced cost functional

$$\mathrm{J}(\boldsymbol{f})=\mathrm{J}(\boldsymbol{f},\mathcal{S}(\boldsymbol{f})),$$

where $\mathcal{S}(\cdot)$ denotes the control-to-state mapping associated with the state system (31). Let $\boldsymbol{h}$ be an admissible perturbation of the control. The corresponding first-order variation of the state is given by

$$\mathfrak{m}=\mathcal{S}'(\boldsymbol{f})\boldsymbol{h}$$

where $\mathcal{S}'(\boldsymbol{f})$ denotes the Gâteaux derivative of the control-to-state mapping. Differentiating the state equation with respect to the control variable shows that $\mathfrak{m}$ satisfies the linearized state equation. Therefore, the solution of the linearized system is precisely the directional derivative of the state with respect to the control. Consequently, the Gâteaux derivative of the reduced cost functional is obtained by applying the chain rule,

$$\mathrm{J}'(\boldsymbol{f})\boldsymbol{h}=\mathrm{J}_{\boldsymbol{v}}(\boldsymbol{f},\boldsymbol{v})\mathfrak{m}+\mathrm{J}_{\boldsymbol{f}}(\boldsymbol{f},\boldsymbol{v})\boldsymbol{h},$$

where $\mathfrak{m}$ is the solution of the linearized state equation.

Remark 7.3. 1. Note that the appearance of the linearized state equation in the expression for the Gâteaux derivative is not coincidental. The linearized equation characterizes the first-order sensitivity of the state with respect to perturbations of the control, and its solution represents the derivative of the control-to-state mapping. Hence, every computation of the derivative of the reduced cost functional necessarily involves the solution of the linearized system.

2. Although the Gâteaux derivative can be expressed in terms of the solution of the linearized state equation, solving the linearized problem for every admissible perturbation $\boldsymbol{h}$ is computationally impractical. To eliminate the dependence on the linearized state, we introduce the adjoint variable. Using the adjoint equation, the terms involving the solution of the linearized system can be transferred to the adjoint variable, yielding an expression for the Gâteaux derivative that depends only on the state, the adjoint state, and the control. The first-order necessary optimality condition is then obtained by requiring the Gâteaux derivative of the reduced cost functional to satisfy the corresponding variational inequality in the presence of control constraints.

Suppose that $(\widetilde{\boldsymbol{f}},\widetilde{\boldsymbol{v}})$ is an optimal control pair. Consider $\boldsymbol{\psi}\in\mathcal{F}_{ad}$ and define

$\boldsymbol{f}_\rho := \widetilde{\boldsymbol{f}} + \rho(\boldsymbol{\psi} - \widetilde{\boldsymbol{f}})$. Thanks to Propositions 5.2 and 5.3, we have

$$\frac{\mathrm{J}(\boldsymbol{f}_\rho, \boldsymbol{v}_\rho) - \mathrm{J}(\widetilde{\boldsymbol{f}}, \widetilde{\boldsymbol{v}})}{\rho}$$

$$= \mathbb{E}\left[\int_0^T \{(\nabla_{\boldsymbol{f}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \boldsymbol{\psi}(t) - \widetilde{\boldsymbol{f}}(t)) + (\nabla_{\boldsymbol{v}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \mathfrak{m}(t))\}\mathrm{d}t\right] + \frac{o(\rho)}{\rho}.$$

Then, the Gâteaux derivative of the cost functional J is given by

$$\lim_{\rho\to 0} \frac{\mathrm{J}(\boldsymbol{f}_\rho, \boldsymbol{v}_\rho) - \mathrm{J}(\widetilde{\boldsymbol{f}}, \widetilde{\boldsymbol{v}})}{\rho}$$

$$= \mathbb{E}\left[\int_0^T \{(\nabla_{\boldsymbol{f}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \boldsymbol{\psi}(t) - \widetilde{\boldsymbol{f}}(t)) + (\nabla_{\boldsymbol{v}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \mathfrak{m}(t))\}\mathrm{d}t\right].$$

Therefore, we have

$$\mathbb{E}\left[\int_0^T \{(\nabla_{\boldsymbol{f}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \boldsymbol{\psi}(t) - \widetilde{\boldsymbol{f}}(t)) + (\nabla_{\boldsymbol{v}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \mathfrak{m}(t))\}\mathrm{d}t\right] \geq 0, \tag{207}$$

where $\mathfrak{m}$ is the unique solution to the linearized problem (129) with $\boldsymbol{\psi}$ replaced by $\boldsymbol{\psi} - \widetilde{\boldsymbol{f}}$.

Let $\tilde{\mathfrak{p}}$ be the unique solution of (191). The application of Proposition 7.1 yields

$$\mathbb{E}\left[\int_0^T (\boldsymbol{\psi}(t) - \widetilde{\boldsymbol{f}}(t), \tilde{\mathfrak{p}}(t))dt\right] = \mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{v}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \mathfrak{m}(t))\mathrm{d}t\right]. \tag{208}$$

Finally, we obtain the following optimality condition, for any $\boldsymbol{\psi} \in \mathcal{F}_{ad}$

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T (\boldsymbol{\psi}(t) - \widetilde{\boldsymbol{f}}(t), \tilde{\mathfrak{p}}(t) + \nabla_{\boldsymbol{f}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)))\mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{v}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \mathfrak{m}(t))\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T (\nabla_{\boldsymbol{f}}\mathfrak{L}(t, \widetilde{\boldsymbol{f}}(t), \widetilde{\boldsymbol{v}}(t)), \boldsymbol{\psi}(t) - \widetilde{\boldsymbol{f}}(t))\mathrm{d}t\right] \\
&\geq 0,
\end{aligned} \tag{209}$$

where we have used (207) and (208). The combination of the preceding sections leads to the proof of Theorem 2.6.

Acknowledgments. We express our deep gratitude to the anonymous referees for the careful reading and valuable suggestions and comments that have improved this manuscript significantly.

Data availability. No data was used for the research described in the article.

Declarations. During the preparation of this work, the authors have not used AI tools.

Conflict of interest. The authors declare no conflict of interest.

Author contributions. All authors contributed equally.

REFERENCES

[1] C. Amrouche and D. Cioranescu, On a class of fluids of grade 3, *Int. J. Non-Linear Mech.*, **32** (1997), 73-88.

[2] J. Babutzka and P. C. Kunstmann, L^q-Helmholtz decomposition on periodic domains and applications to Navier-Stokes equations, *J. Math. Fluid Mech.*, **20** (2018), 1093-1121.

[3] H. I. Breckner, *Approximation and Optimal Control of the Stochastic Navier–Stokes Equation*, Halle (Saale), (Ph.D. Thesis), 1999.

[4] A. V. Busuioc, The regularity of bidimensional solutions of the third grade fluids equations, *Math. Comput. Modelling*, **35** (2002), 733-742.

[5] A. V. Busuioc and D. Iftimie, A non-Newtonian fluid with Navier boundary conditions, *J. Dyn. Differ. Equ.*, **18** (2006), 357-379.

[6] V. Busuioc and D. Iftimie, Global existence and uniqueness of solutions for the equations of third grade fluids, *Int. J. Non-Linear Mech.*, **39** (2004), 1-12.

[7] N. Chemetov and F. Cipriano, Optimal control for two-dimensional stochastic second grade fluids, *Stoch. Processes Appl.*, **128** (2018), 2710-2749.

[8] N. V. Chemetov and F. Cipriano, Optimal control of Newtonian fluids in a stochastic environment, *SIAM J. Math. Anal.*, **57** (2025), 364-403.

[9] J.-Y. Chemin and C.-J. Xu, Inclusions de Sobolev en calcul de Weyl-Hörmander et champs de vecteurs sous-elliptiques, *Ann. Sci. École Norm. Sup. (4)*, **30** (1997), 719-751.

[10] F. Cipriano, P. Didier and S. Guerra, Well-posedness of stochastic third grade fluid equation, *J. Differ. Equ.*, **285** (2021), 496-535.

[11] N. J. Cutland and K. Grzesiak, Optimal control for two-dimensional stochastic Navier-Stokes equations, *Appl. Math. Optim.*, **55** (2007), 61-91.

[12] G. da Prato and A. Debussche, Dynamic programming for the stochastic Navier-Stokes equations, *Math. Model. Numer. Anal.*, **34** (2000), 459-475.

[13] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions, 2nd ed.*, Cambridge University Press, 2014.

[14] R. L. Fosdick and K. R. Rajagopal, Thermodynamics and stability of fluids of third grade, *Proc. Roy. Soc. London Ser. A*, **369** (1980), 351-377.

[15] L. Grafakos, *Classical Fourier Analysis*, Third edition, Graduate Texts in Mathematics, Springer, New York, 2014.

[16] T. Hayat, I. Ullah, A. Alsaedi and A. Bashir, Impact of temperature dependent heat source and non-linear radiative flow of third grade fluid with chemical aspects, *Therm Sci.*, **24** (2020), 1173-1182.

[17] W. Liu and M. Röckner, *Stochastic Partial Differential Equations: An Introduction*, Springer, Berlin, 2015.

[18] J. L. Menaldi and S. S. Sritharan, Impulse control of stochastic Navier-Stokes equations, *Nonlinear Anal.*, **52** (2003), 357-381.

[19] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Second edition, McGraw-Hill Book Co., New York, 1984.

[20] M. Parida and S. Padhy, Electro-osmotic flow of a third-grade fluid past a channel having stretching walls, *Nonlinear Eng.*, **8** (2019), 56-64.

[21] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, 44, Springer-Verlag, New York, 1983.

[22] G. J. Reddy, A. Hiremath and M. Kumar, Computational modeling of unsteady third-grade fluid flow over a vertical cylinder: A study of heat transfer visualization, *Results Phys.*, **8** (2018), 671-682,

[23] R. S. Rivlin and J. L. Ericksen, Stress-deformation relations for isotropic materials, *Arch. Rational Mech. Anal.*, **4** (1995), 323-425.

[24] J. C. Robinson, *Infinite-Dimensional Dynamical Systems: An Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors*, Cambridge University Press, 2001.

[25] J. C. Robinson, J. L. Rodrigo and W. Sadowski, *The Three-Dimensional Navier–Stokes Equations, Classical Theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, UK, 2016.

[26] A. Sequeira and J. Videman, Global existence of classical solutions for the equations of third grade fluids, *J. Math. Phys. Sci.*, **29** (1995), 47-69.

[27] S. S. Sritharan, Deterministic and stochastic control of Navier–Stokes equation with linear, monotone, and hyperviscosities, *Appl. Math. Optim.*, **41** (2000), 255-308.

[28] Y. Tahraoui and F. Cipriano, Optimal control of two dimensional third grade fluids, *J. Math. Anal. Appl.*, **523** (2023), Paper No. 127032, 33 pp.

[29] Y. Tahraoui and F. Cipriano, Optimal control of third grade fluids with multiplicative noise, *ESAIM Control Optim. Calc. Var.*, **31** (2025), Paper No. 16.

[30] Y. Tahraoui and F. Cipriano, Local strong solutions to the stochastic third grade fluid equations with Navier boundary conditions, *Stoch. PDE: Anal. Comp.*, **12** (2024), 1699-1744.

[31] R. Temam, *Navier-Stokes Equations, Theory and Numerical Analysis*, North-Holland, Amsterdam, 1984.
[32] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer, New York, 1997.

Received March 9, 2026; revised July 5, 2026.